

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: September 8, 2005, 11:52:12 ; Search time 189.5 Seconds  
(without alignments)  
4702.353 Million cell updates/sec

Title: US-10-723-180-1  
Perfect score: 2181  
Sequence: 1 gggggcaggcaatgagagctg.....attactcagtctaacaata 1152

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delgap 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p model -DEV=xlP  
-Q=/cgn2\_1/USFTO\_spool\_p/US10723180/runat\_08092005\_123255\_29226/app\_query.fasta\_1.1351  
-DB=A\_Geneseq -QMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=us10723180 @CNC 1.1 224 @runat\_08092005\_123255\_29226 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A\_Geneseq 16Dec04:.\*  
1: Geneseqp19808:.\*  
2: Geneseqp19908:.\*  
3: Geneseqp20008:.\*  
4: Geneseqp20018:.\*  
5: Geneseqp20028:.\*  
6: Geneseqp20038a:.\*  
7: Geneseqp20038b:.\*  
8: Geneseqp20048:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description
1	1685	77.3	323 6 ABUS56649 Lung canc
2	1685	77.3	323 7 ADN38936 Cancer/an
3	1685	77.3	323 8 ADS13763 Human aqu
4	1458.5	66.9	300 5 ABBS7089 Mouse isc
5	617	28.3	265 5 Abg96280 Human ova
6	617	28.3	265 5 ABBS97500 Novel hum
7	617	28.3	265 6 ABBS2054 Human cer
8	617	28.3	265 7 ADD47950 Human Pro
9	617	28.3	265 7 ADE62232 Human Pro
10	617	28.3	265 8 ADJ75631 Marker ge

11	617	28.3	265	8	ADJ75701	Marker ge
12	617	28.3	265	8	ADL70563	Cervical
13	614	28.2	265	2	Aaw55787	Rat aquap
14	614	28.2	265	2	Aaw94319	Rat aquap
15	614	28.2	265	7	ADD47948	Rat Prote
16	614	28.2	265	8	ADJ76346	Marker ge
17	614	28.2	265	8	ADJ76388	Marker ge
18	614	28.2	263	2	AAR25424	Human MIP
19	612	28.1	263	2	ADD48472	Human Pro
20	612	28.1	263	7	ADBS63588	Human Pro
21	612	28.1	118	3	AAG02008	Human sec
22	603	27.6	282	6	ABR41500	Human DIT
23	600.5	27.5	261	7	ADD48470	Rat Prote
24	600	27.5	261	7	ADBS63586	Rat Prote
25	600	27.5	261	7	ADBS13759	Human aqu
26	597	27.4	271	8	ADBS13759	Human aqu
27	594	27.2	271	2	AAR51070	A water c
28	589	27.0	271	7	ADG32000	Human hom
29	561	25.7	269	2	Aaw55786	Human aqu
30	561	25.7	269	2	Aaw94322	Human aqu
31	561	25.7	269	6	ABP58050	Human aqu
32	561	25.7	269	7	ADBS5200	Human aqu
33	561	25.7	269	8	ABMS0877	Tumour-as
34	561	25.7	269	8	ADBS13757	Human aqu
35	525.5	24.1	222	8	ADL70567	Cervical
36	508	23.3	282	7	ADK50990	Human NOV
37	496.5	22.8	239	4	ABBS64172	Drosophil
38	447	20.5	72	4	ABBS29884	Peptide #
39	447	20.5	72	4	ABBS20480	Protein #
40	447	20.5	72	4	AAM68252	Human bon
41	447	20.5	72	4	AAM55880	Human bra
42	447	20.5	72	4	ABG49906	Human liv
43	447	20.5	72	5	ABG37787	Human pep
44	443.5	20.3	253	3	AAG48889	Arabidops
45	443.5	20.3	278	3	AAG48888	Arabidops

ALIGNMENTS

RESULT 1  
ABUS56649  
ID ABUS56649 standard; protein; 323 AA.  
XX

AC ABUS56649;  
XX  
DT 02-APR-2003 (first entry)  
XX  
DE Lung cancer-associated polypeptide #242.  
XX

KW Lung cancer-associated polypeptide; cytostatic; emphysema;  
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

OS Unidentified.

PN WO200286443-A2.

PD 31-OCT-2002.

PF 18-APR-2002; 2002WO-US012476.

PR 18-APR-2001; 2001US-0284770P.

PR 10-MAY-2001; 2001US-0290492P.

PR 09-NOV-2001; 2001US-0339245P.

PR 13-NOV-2001; 2001US-0350666P.

PR 29-NOV-2001; 2001US-034370P.

PR 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

PI Aziz N, Murray R;

XX WPI: 2003-093161/08.  
 DR N-PSDB; ABX76378.  
 XX  
 PT Detecting a lung cancer-associated transcript in a cell from a patient  
 PT for treating lung cancer, by contacting a biological sample from the  
 PT patient with a polynucleotide that exhibits increased or decreased  
 XX expression in lung cancer.  
 XX  
 PS Claim 27; Page 375; 453pp; English.  
 XX  
 CC The invention relates to a method for detecting a lung cancer-associated  
 CC transcript in a cell from a patient, comprising contacting a biological  
 CC sample from the patient with a polynucleotide that selectively hybridizes  
 CC to a sequence that is at least 80 % identical to a gene that exhibits  
 CC increased or decreased expression in lung cancer samples. Lung cancer-  
 CC associated polynucleotides and polypeptides are used for identifying a  
 CC compound that modulates a lung cancer-associated polypeptide, for  
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung  
 CC cancer in a patient and for treating a mammal having lung cancer by  
 CC administering a modulatory compound identified. The methods are useful  
 CC for treating lung cancer, such as small cell lung cancer, non-small cell  
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,  
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,  
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and  
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful  
 CC for diagnostic purposes and as targets for screening for therapeutic  
 CC compounds that modulate lung cancer, such as antibodies. Sequences  
 CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the  
 CC invention  
 XX  
 SQ Sequence 323 AA;

## Alignment Scores:

Pred. No.: 2,09e-158 Length: 323  
 Score: 1685.00 Matches: 323  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 77.26% Indels: 0  
 DB: 6 Gaps: 0

US-10-723-180-1 (1-1152) x ABUS6649 (1-323)

QY 40 ATGAGTGACAGACCCACAGACGCGGTGGGTGAAGTGGACCTTGTGTACGAGAG 99  
 DB 1 MetSerAspArgProThrAlaArgArgTrpGlyLysCysGlyProLeuCysThrArgGlu 20  
 QY 100 AACATCATGGTGGCTTCAAGGGGTCTGGACCTTCTGGAAGCAGTCAAGGG 159  
 DB 21 AsnIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAla 40  
 QY 160 GAATTTCTGGCATGCTTATTTTGTCTCTCAGCGTGGATCCACCACTCACTGGGT 219  
 DB 41 GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGly 60  
 QY 220 GGAACAGAAAGCCTTACCGTTCGACATGGTCTCATCTCCCTTGGCTTGGACTCAGC 279  
 DB 61 GlyThrGluLysProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer 80  
 QY 280 ATTGCAACCATGGTGCAGTGTCTTGGCCATATCAGCGGTGGCCACATCAACCTCGAGTG 339  
 DB 81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100  
 QY 340 ACTGTGGCCATGGTGTGACAGGAGATCAGCATCGCAAGTCTGTCTTCTACATCGCA 399  
 DB 101 ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTrpIleAla 120  
 QY 400 GCCCAGTGCCTGGGGCCATCATTCGAGCAGGAACTCTATCTGTGTCACTCCCACT 459  
 DB 121 AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTrpValThrProSer 140  
 QY 460 GTGGTGGGAGGCGTGGGAGTCCACCATGGTTCATGGAAATCTTACCGCTGGTGTCTC 519

DB 141 ValValGlyGlyLeuGlyValThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeu 160  
 QY 520 CTGGTTAGTTGATAATCACAATTCATTTCAATGGTGTACTATCTTTGCCAGCTGTGATCC 579  
 DB 161 LeuValGluLeuIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180  
 QY 580 AAAAGCAGCTGATGCACCTGCTCAATAGCTTTAGCAATTTGATTTCTGTGCAATTCGA 639  
 DB 181 LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGly 200  
 QY 640 CATTTATTGGCAATCAATTATATCTGTCAGCATGAATCCGCCGATCCTTTTGGACCT 699  
 DB 201 HisLeuPheAlaIleAsnTrpThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220  
 QY 700 CGAGTTATCATGGGAAATTCGGAACCAATTCGATATATTCGGTTGGCCCATCATAGA 759  
 DB 221 AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTrpValGlyProIleIleGly 240  
 QY 760 GCTGTCTCTGCTGGTGGCTTTATGAGTATCTCTCTGTCAGATGTTGAATTCAAACGT 819  
 DB 241 AlaValLeuAlaGlyGlyLeuTrpGluTrpValPheCysProAspValGluPheLysArg 260  
 QY 820 CGTTTAAAGAAAGCCTTCAGCAAGCTGCCAGCAAAACAAAGGAAGCTACATGAGGTG 879  
 DB 261 ArgPheLysGluAlaPheSerLysAlaAlaGlnThrLysGlySerTrpMetGluVal 280  
 QY 880 GAGGACAAACGAGGTTCAGGTAGACGATGACCTGATTCCTTAAACCTTGGAGTGGTGCAT 939  
 DB 281 GluAspAsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGlyValValHis 300  
 QY 940 GTGATTGACGTTGACCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 999  
 DB 301 ValIleAspValAspArgGlyGluGluLysLysGlyLysAspGlnSerGlyGluValLeu 320  
 QY 1000 TCTTCAGTA 1008  
 DB 321 SerSerVal 323  
 RESULT 2  
 ID ADN38936  
 XX ADN38936 standard; protein; 323 AA.  
 XX AC ADN38936;  
 XX DT 17-JUN-2004 (first entry)  
 XX DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:254.  
 XX KW Human; differential expression; cancer; angiogenic disorder;  
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;  
 KW inflammatory disease; autoimmune disease;  
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;  
 KW detection; diagnosis; prognosis; drug screening; drug targeting;  
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;  
 KW vulnery; gene therapy; vaccine.  
 XX OS Homo sapiens.  
 XX PN WO2003042661-A2.  
 XX PD 22-MAY-2003.  
 XX PF 13-NOV-2002; 2002WO-US036810.  
 XX PR 13-NOV-2001; 2001US-0350666P.  
 PR 21-NOV-2001; 2001US-0332464P.  
 PR 29-NOV-2001; 2001US-0334393P.  
 PR 03-DEC-2001; 2001US-0335394P.  
 PR 14-DEC-2001; 2001US-0340376P.  
 PR 08-JAN-2002; 2002US-0347211P.  
 PR 10-JAN-2002; 2002US-0347349P.  
 PR 08-FEB-2002; 2002US-0355250P.  
 PR 13-FEB-2002; 2002US-0356714P.

PR 20-FEB-2002; 2002US-0359077P.  
 PR 29-MAR-2002; 2002US-0368809P.  
 PR 04-APR-2002; 2002US-0370110P.  
 PR 12-APR-2002; 2002US-0372246P.  
 PR 05-JUN-2002; 2002US-0386614P.  
 PR 16-JUL-2002; 2002US-0396839P.  
 PR 22-JUL-2002; 2002US-0397775P.  
 PR 22-JUL-2002; 2002US-0397845P.  
 PR 09-SEP-2002; 2002US-0409450P.  
 XX (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;  
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;  
 XX WPI; 2003-468649/44.  
 DR N-PSDB; ADN38935.  
 XX  
 XX Determining the presence or absence of a pathological cell in a patient,  
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting  
 PT a nucleic acid in a biological sample.  
 XX  
 XX Claim 12; SEQ ID NO 254; 1385pp; English.  
 XX  
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)  
 CC whose expression is upregulated or downregulated in specific cancers or  
 CC other diseases such as angiogenic or fibrotic disorders, and to methods  
 CC of determining the presence or absence of a pathological cell in a  
 CC patient by detecting a nucleic acid at least 80% identical to those of  
 CC the invention or by detecting a polypeptide of the invention. The  
 CC invention also relates to expression vectors and host cells comprising a  
 CC nucleic acid of the invention; antibodies which specifically bind a  
 CC polypeptide of the invention; use of such antibodies for drug targeting;  
 CC and methods of screening for modulators of activity or expression of the  
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides, and  
 CC antibodies and methods are useful for diagnosing, prognosing and treating  
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,  
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal  
 CC neovascularisation syndromes, scarring and uterine fibroids. They may  
 CC also be useful in wound healing and in contraception. The present  
 CC sequence represents a polypeptide of the invention.  
 XX  
 XX Sequence 323 AA;  
 SQ

Alignment Scores:  
 Pred. No.: 2,09e-158 Length: 323  
 Score: 1685.00 Matches: 323  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 77.26% Indels: 0  
 DB: 7 Gaps: 0

US-10-723-180-1 (1-1152) x ADN38936 (1-323)  
 QY 40 ATGAGTGACAGACCCACAGACGGCGTGGGTGAGTGTGACCTTTGTGTACACGAGAG 99  
 Db 1 MetSerAspArgProThrAlaArgArgTgPglyLysCysGlyProLeuCysThrArgGlu 20  
 QY 100 AACATCATGGTGGCTTCAAGGGGTCTGAGCTCAAGCTTCTGGAAACAGTCAAGCG 159  
 Db 21 AnileMetValAlaPheLysGlyValTgPthrGlnAlaPheTrpLysAlaValThrAla 40  
 QY 160 GAATTTCTGGCCATGCTTATTTTGTCTCTCAGCTGGATCCACCATCAACTCACTGGGT 219  
 Db 41 GluPheLeuAlaMetLeullePheValLeuLeuSerLeuGlySerThrIleAsnTrpGly 60  
 QY 220 GGAACAGAAAGCCTTTACCGGTGACATGGTGTCTCATCTCCCTTTGCTTGGACTCAGC 279  
 Db 61 GlyThrGluLysProLeuProValAspMetValLeulleSerLeuCysPheGlyLeuSer 80  
 QY 280 ATTGCAACCATGGTGCAGTGTCTTGGCCATATCAGCGGTGGCCACATCAACCTCGAGTG 339  
 Db 81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyHisIleAsnProAlaVal 100

QY 340 ACTGTGGCCATGGTGTGACACGAGGATCAGCATCGCCCAAGTCTGTCTTCTACATCGCA 399  
 Db 101 ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrlleAla 120  
 QY 400 GCCAGTGGCTGGGGCCCATCATTTGAGCAGAGAACTCTATCTATGTGTACACCTCCAGT 459  
 Db 121 AlaGlnCysLeuGlyAlaIlelleGlyAlaGlylleLeuTyrlleValThrProSer 140  
 QY 460 GTGGTGGGAGGCTGGGAGTCAACATGGTTCAATGAAATCTTACCGCTGGTATGCTCTC 519  
 Db 141 ValValGlyGlyLeuGlyValThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeu 160  
 QY 520 CTGTTGTAGTTGATTAATCACATTTCAATTTGGTGTGTACTATCTTTGCCAGCTGTGATTC 579  
 Db 161 LeuValGluLeullelleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180  
 QY 580 AAACGGACTGATGTCACTGGCTCAATAGCTTTAGCAATGGATTTTCTTTGCAATGGA 639  
 Db 181 LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGly 200  
 QY 640 CATTTATTGCAATCAATTTATCTGTCGACGATGAATCCCGCCGATCCCTTTGGACCT 699  
 Db 201 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220  
 QY 700 CGAGTTATCATGGGAAATTTGGGAAACCATTTGGATATATTGGTGTGGGCCCATCATAGGA 759  
 Db 221 AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIlelleGly 240  
 QY 760 GCTGCTCTGCTGGTGGCTTTATGAGTATGTCCTTCTGTCCAGATGTTGAATCAAACT 819  
 Db 241 AlaValLeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluPheLysArg 260  
 QY 820 CGTTTAAAGAGCCCTTCAAGCAAGCTGCCAGCAAGCAAGAGGAGCTACATGAGGCTG 879  
 Db 261 ArgPheLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluVal 280  
 QY 880 GAGGACAAACAGGAGTCAGGTAGAGACGATGACCTGATTTCTAAAACTCGAGTGGTGCAT 939  
 Db 281 GluAspAsnArgSerGlnValGluThrAspPheLeulleLeuLysProGlyValValHis 300  
 QY 940 GTGATTGACGTTGACCGGGGAGAGAGAGAAAGGGGAAAGCAATCTGGAGAGGTATTG 999  
 Db 301 ValIleAspValAspArgGlyGluGlyLysGlyLysAspGlnSerGlyGluValLeu 320  
 QY 1000 TCTTTCAGTA 1008  
 Db 321 SerSerVal 323  
 RESULT 3  
 ADS13763  
 ID ADS13763 standard; protein; 323 AA.  
 XX  
 AC ADS13763;  
 DT  
 XX 02-DEC-2004 (first entry)  
 DE Human aquaporin, AQP4.  
 XX  
 KW Aquaporin; water channel; hypotensive; cardiovascular; hepatotropic;  
 KW nephrotropic; endocrine; cytostatic; antiallergic; antibacterial;  
 KW immunosuppressive; virucide; fungicide; human; AQP4.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004077010-A2.  
 XX  
 PD 10-SEP-2004.  
 XX  
 PF 03-FEB-2004; 2004WO-US003022.  
 XX  
 PF 03-FEB-2003; 2003US-0444846P.  
 XX  
 XX

```
(Scri ) SCRIPPS RES INST.
(DANI/) DANIELS M.
(YEAG/) YEAGER M.
XX
XX Daniels M, Yeager M;
XX
XX WPI; 2004-662063/64.
XX N-PSDB; ADS13764.
XX GENBANK; AAH22286.
XX
XX Identifying agents that can modulate water channel activity, useful for
XX PT treating hypertension or congestive heart failure and in affecting the
XX PT hydraulic conductivity of other plant tissues, comprises obtaining an
XX PT aquaporin-expressing cell.
XX
XX Claim 3; SEQ ID NO 7; 69pp; English.
XX
XX The invention relates to identifying agents that can modulate water
XX channel activity. The method involves obtaining an aquaporin-expressing
XX cell that is osmotically sensitive in an osmotically acceptable solution;
XX contacting the aquaporin-expressing cell that is osmotically sensitive
XX with a test agent; osmotically shocking the aquaporin-expressing cell
XX that is osmotically sensitive in the presence of the test agent; and
XX observing whether water channel activity of the aquaporin is modulated.
XX The identified agents are useful for treating hypertension, congestive
XX heart failure and clinical states relating to abnormal water balance,
XX e.g. hepatic cirrhosis, nephrotic syndrome, renal failure, endocrine
XX disorders, cancers, allergic reactions, or sepsis from viral, bacterial,
XX or fungal organism, in altering water uptake by plant roots, and in
XX affecting the hydraulic conductivity of other plant tissues. The present
XX sequence represents a human aquaporin, AQP4 polypeptide.
XX
XX Sequence 323 AA;
SQ
Alignment Scores:
Pred. No.: 2,09e-158 Length: 323
Score: 1685.00 Matches: 323
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 77.26% Indels: 0
DB: 8 Gaps: 0
US-10-723-180-1 (1-1152) x ADS13763 (1-323)
QY 40 ATGAGTCAGACCCACAGCAGCGGTGGTAAAGTGTGGACCTTTGTGTACACAGAG 99
Db 1 MetSerAspArgProThrAlaArgTrpGlyLysCysArgProLeuCysThrArgGlu 20
QY 100 AACATCATGGTGGCTTCAAGGGCTCGACTCAAGCTTTCTGGAAAGCAGTCACGG 159
Db 21 AsnIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAla 40
QY 160 GAATTTCTGGCCATGCTATTATTTGTTCCTCCAGCTGGGATCCACCATCAACTGGGT 219
Db 41 GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGly 60
QY 220 GGAACAGAAAAGCCTTTACCGTCGACATGGTTCTCATCTCCCTTGTCTTGCATCAGC 279
Db 61 GlyThrGluLysProLeuProValAspMetValLeuIleSerLeuLysPheGlyLeuSer 80
QY 280 ATTGCAACCATGGTCAGTGTCTTGGCCATATCAGGGTGGCCACATCAACCTCCAGTG 339
Db 81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100
QY 340 ACTGTGGCCATGGTGTGACACAGGAGATCAGCATCGCCCAAGTCTGTCTTCTACATCGCA 399
Db 101 ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTrpIleAla 120
QY 400 GCCCAGTGCCTGGGGCCCATCATTTGGAGCAGGAATCTCTATCTGTGTGCACCTCCCACT 459
Db 121 AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyLeuValThrProSer 140
QY 460 GTGGTGGGAGGCTGGGAGTCACCATGGTTCATGTGGAATCTTACCGCTGGTCAATGGTCTC 519
Db
PA (Scri ) SCRIPPS RES INST.
PA (DANI/) DANIELS M.
PA (YEAG/) YEAGER M.
XX
XX Daniels M, Yeager M;
XX
XX WPI; 2004-662063/64.
XX N-PSDB; ADS13764.
XX GENBANK; AAH22286.
XX
XX Identifying agents that can modulate water channel activity, useful for
XX PT treating hypertension or congestive heart failure and in affecting the
XX PT hydraulic conductivity of other plant tissues, comprises obtaining an
XX PT aquaporin-expressing cell.
XX
XX Claim 3; SEQ ID NO 7; 69pp; English.
XX
XX The invention relates to identifying agents that can modulate water
XX channel activity. The method involves obtaining an aquaporin-expressing
XX cell that is osmotically sensitive in an osmotically acceptable solution;
XX contacting the aquaporin-expressing cell that is osmotically sensitive
XX with a test agent; osmotically shocking the aquaporin-expressing cell
XX that is osmotically sensitive in the presence of the test agent; and
XX observing whether water channel activity of the aquaporin is modulated.
XX The identified agents are useful for treating hypertension, congestive
XX heart failure and clinical states relating to abnormal water balance,
XX e.g. hepatic cirrhosis, nephrotic syndrome, renal failure, endocrine
XX disorders, cancers, allergic reactions, or sepsis from viral, bacterial,
XX or fungal organism, in altering water uptake by plant roots, and in
XX affecting the hydraulic conductivity of other plant tissues. The present
XX sequence represents a human aquaporin, AQP4 polypeptide.
XX
XX Sequence 323 AA;
SQ
Alignment Scores:
Pred. No.: 2,09e-158 Length: 323
Score: 1685.00 Matches: 323
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 77.26% Indels: 0
DB: 8 Gaps: 0
US-10-723-180-1 (1-1152) x ADS13763 (1-323)
QY 40 ATGAGTCAGACCCACAGCAGCGGTGGTAAAGTGTGGACCTTTGTGTACACAGAG 99
Db 1 MetSerAspArgProThrAlaArgTrpGlyLysCysArgProLeuCysThrArgGlu 20
QY 100 AACATCATGGTGGCTTCAAGGGCTCGACTCAAGCTTTCTGGAAAGCAGTCACGG 159
Db 21 AsnIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAla 40
QY 160 GAATTTCTGGCCATGCTATTATTTGTTCCTCCAGCTGGGATCCACCATCAACTGGGT 219
Db 41 GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGly 60
QY 220 GGAACAGAAAAGCCTTTACCGTCGACATGGTTCTCATCTCCCTTGTCTTGCATCAGC 279
Db 61 GlyThrGluLysProLeuProValAspMetValLeuIleSerLeuLysPheGlyLeuSer 80
QY 280 ATTGCAACCATGGTCAGTGTCTTGGCCATATCAGGGTGGCCACATCAACCTCCAGTG 339
Db 81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100
QY 340 ACTGTGGCCATGGTGTGACACAGGAGATCAGCATCGCCCAAGTCTGTCTTCTACATCGCA 399
Db 101 ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTrpIleAla 120
QY 400 GCCCAGTGCCTGGGGCCCATCATTTGGAGCAGGAATCTCTATCTGTGTGCACCTCCCACT 459
Db 121 AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyLeuValThrProSer 140
QY 460 GTGGTGGGAGGCTGGGAGTCACCATGGTTCATGTGGAATCTTACCGCTGGTCAATGGTCTC 519
Db
141 ValValGlyGlyLeuGlyValThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeu 160
520 CTGGTTGAGTTGATAATCACATTTCAATTTGGTGTCTTACTATCTTTCAGCTGTGATTC 579
161 LeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180
580 AACGGAGCTGATGCTACCTGCTCAATAGCTTTAGCAATTTGGAATTTCTGTTGCAATTGGA 639
181 LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGly 200
640 CATTTATTTGCAATCAATTTACTGTCGCCAGCATGAATCCGCCCGCATCTTTTGGACCT 699
201 HisLeuPheAlaIleAsnTyThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220
700 GCAGTTATCATGGGAAATTTGGGAAACCATTTGGATATATTTGGTGTGGCCCATCATAGA 759
221 AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyTrpValGlyProIleIleGly 240
760 GCTGTCTCTGCTGGTGGCTTTATGAGTATCTTCTCTGTCAGATGTTGAATTCAAACGT 819
241 AlaValLeuAlaGlyGlyLeuTyArgValPheCysProAspValGluPheLysArg 260
820 CGTTTAAAGAGCTTTCAGCAAGCTGCCAGCAAGCTGCCAGCAAGCTGCCAGCTACATGAGGTG 879
261 ArgPheLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyMetGluVal 280
880 GAGACACACAGAGTTCAGGTAGAGACGATGATGATGATGATGATGATGATGATGATGATGAT 939
281 GluAspAsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGlyValValHis 300
940 GTGATTGACGTTCGCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 999
301 ValIleAspValAspArgGlyGluGluLysLysGlyLysAspGlnSerGlyGluValLeu 320
1000 TCTTCAGTA 1008
321 SerSerVal 323
RESULT 4
ABB57089
ID ABB57089 standard; protein; 300 AA.
XX
AC ABB57089;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related protein sequence SEQ ID NO:194.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP004192.
XX
PR 18-MAY-2000; 2000JP-00145977.
XX
PA (UUNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI; 2002-034733/04.
DR N-PSDB; ABI199325.
XX
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
XX PT expression levels of particular genes defined in the specification or by
XX PT determining the expression profile of a gene group comprising these
XX PT genes.
```







Score:	617.00	Matches:	132
Percent Similarity:	54.64%	Conservative:	38
Best Local Similarity:	50.19%	Mismatches:	81
Query Match:	28.29%	Indels:	12
DB:	5	Gaps:	4
US-10-723-180-1 (1-1152) x ABR97500 (1-265)			
Qy	118	AAAGGGGTCTGGACTCAAGCTTTCTGGAAAGCAGTCCACAGCGGAATTTCTGGCCATGCTT	177
Db	3	LysGluValCysSerValAlaPheLeuLeuAlaValPheAlaGluPheLeuAlaThrLeu	22
Qy	178	ATTTTGTCTCTCAGCCTGGGATCCACCATCACTGCGGTGGAACAGAAAGCCTTTA	237
Db	23	IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp-----ProSerAlaLeu	39
Qy	238	CCGGTCGACATGTTCTCATCTCCCTTCTTGGACTCAGCATTCGACCATCCACCATGCTGCAG	297
Db	40	ProThr---IleLeuGlnIleAlaLeuAlaPheGlyLeuAlaIleGlyThrLeuAlaGln	58
Qy	298	TGCTTTGGCCATATCAGCGTGGCCACATCAACCTGCGAGTGAATCCCTGCGAGTGTGTC	357
Db	59	AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuVal	78
Qy	358	ACAGGAAGATCAGCATCGCAGTCTGTCTTACATCGCAGCCGAGTCTGGGGGCC	417
Db	79	GlyAsnGlnIleSerLeuLeuArgAlaPhePheThrValAlaAlaGlnLeuValGlyAla	98
Qy	418	ATCATTGGAGCAAACTCTATCTGGTCACACCTCCAGTGTGTGGAGCGCTGGGA	477
Db	99	IleAlaGlyAlaGlyIleLeuThrGlyValAlaProLeuAsnAlaAlaArgGlyAsnLeuAla	118
Qy	478	GTCACCATGTTTCATGGAAATCTTACCGCTGGTCTCTCTGCTGAGTTGATAATC	537
Db	119	ValAsnAlaLeuAsnAsnThrThrGlnGlyAlaMetValValGluLeuIleLeu	138
Qy	538	ACATTTCAATGGTGTACTATCTTTGCCAGCTGTGATTCACAAAGGACTGATGTCAT	597
Db	139	ThrPheGlnLeuAlaLeuCysIlePheAlaSerThrAspSerArgArgThrSerProVal	158
Qy	598	GGCTCAATAGCTTAGCAATTTCTGTTGCAATGGACATTTATTTCATCAT	657
Db	159	GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleThr	178
Qy	658	TATACTGGTCCAGCATGAATCCCGCCGATCTTTGGACCTCAGTTATCTGGGAAT	717
Db	179	PheThrGlyCysSerMetAsnProAlaAlaArgSerPheGlyProAlaValValMetAsnArg	198
Qy	718	TGGGAA---AACATTGGATATATGGTTGGGCCCATCATAGGACTCTCTCGCTGT	774
Db	199	PheSerProAlaHisTrpValPheTrpValGlyProIleValGlyAlaValLeuAlaAla	218
Qy	775	GGCTTTATGAGTATGCTCTCTCCAGATGTTGAATTCAAAGCTGCTTTAAAGAGCC	834
Db	219	IleLeuThrPheThrLeuLeuPheProAsn-----SerLeuSer	231
Qy	835	TTACGAAAGCTGCCAGCAACAAAGGAAGCTACATGGAGTGGAGGACACAGGAGT	894
Db	232	LeuSerGluArgValAlaIleIleLeuGlyThrTyGluProAspGluAspTrpGluGlu	251
Qy	895	CAGGTAGAG	903
Db	252	GlnArgGlu	254
RESULT 7			
ABR92054			
ID	ABR92054 standard; protein; 265 AA.		
XX			
XX			
XX	ABR92054;		
XX			
DT	10-SEP-2003 (first entry)		
XX			
DE	Human cervical cancer cell marker protein SEQ ID NO:16.		



QY 478 GTCCACCATGTTTCATGGAATCTTACCGCTGGTGCATGGTCTCCTGGTTGAGTTGATAATC 537  
Db 119 ValAsnAlaLeuAsnAsnThrThrGlnGlyGlnAlaMetValValGluLeuLeuLeu 138  
QY 538 ACATTTCAATGCTGTTTACTATCTTTGCCAGCTGTGATTCCAAACGGAGTGCATCTCACT 597  
Db 139 ThrPheGlnLeuAlaLeuCysIlePheAlaSerThrAspSerArgThrSerProVal 158  
QY 598 GGCCTCAATAGCTTTAGCAATTTGATTTCTGTTGGCAATTTGGACATTTATTTGCAATCAAT 657  
Db 159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178  
QY 658 TATACCTGGCCAGCATGAATCCCGCCGATCTTTGGACCTCCTGACGTTATCATGGGAAT 717  
Db 179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValValMetAsnArg 198  
QY 718 TGGGAA--AACCATTGGATATATGCTTCTGTCAGATGTTGAATTCAAACGTCGTTTAAAGAGCC 834  
Db 199 PheSerProAlaHisTrpValPheTrpValGlyProIleValGlyAlaValLeuAlaAla 218  
QY 775 GGCCTTTAGTATGCTTCTGTCAGATGTTGAATTCAAACGTCGTTTAAAGAGCC 834  
Db 219 IleLeuTyrPheTyrLeuLeuPheProAsn-----SerLeuSer 231  
QY 835 TTCAGCAAGCTGCCAGCAAAACAAAGGAGCTACATGAGGTGGAGGACCAACAGGAGT 894  
Db 232 LeuSerGluArgValAlaIleIleLysGlyThrTyrGluProAspGluAspTrpGluGlu 251  
QY 895 CAGTAGAG 903  
Db 252 GlnArgGlu 254

## RESULT 9

ADE62232

ID ADE62232 standard; protein; 265 AA.

AC ADE62232;

DT 29-JAN-2004 (first entry)

DE Human Protein NP\_001642, SEQ ID NO 8161.

KW Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

OS Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GENO ) GEN HOSPITAL CORP.

XX (FARB ) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI: 2003-268312/26.

XX GENBANK; NP\_001642.

XX New composition comprising two or more isolated polypeptides, useful for

XX preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat

XX or human polynucleotides or a polynucleotide which represents a fragment,

CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 265 AA;

## Alignment Scores:

Pred. NO.:	4,276-52	Length:	265
Score:	617.00	Matches:	132
Percent Similarity:	64.64%	Conservative:	38
Best Local Similarity:	50.19%	Mismatches:	81
Query Match:	28.23%	Indels:	12
DB:	7	Gaps:	4

US-10-723-180-1 (1-1152) x ADE62232 (1-265)

QY 118 AAAGGGTCTGGACTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTT 177

Db 3 LysGluValCysSerValAlaPheLeuLysAlaValPheAlaGluPheLeuAlaThrLeu 22

QY 178 ATTTTGTCTCTCCTCAGCTGGGATCCACCATCAACTGGGGTGAACAGAAAAGCCTTTA 237

Db 23 IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp-----ProSerAlaLeu 39

QY 238 CCGGTGCGACATGGTTCATCTCCCTTTGCTTTGGACTCAGCATTCGAAACATGTTGCGAG 297

Db 40 ProThr---IleLeuGlnIleAlaLeuAlaPheGlyLeuAlaIleGlyThrLeuAlaGln 58

QY 298 TGCTTTGGCCATATCAGCGGTGGCCACATCAACCTCGAGTCAGTCTGGCCATGGTGTGC 357

Db 59 AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuVal 78

QY 358 ACCAGGAAGATCAGCATCGCCAAAGTCTGTCTTCTACATCGCAGCCAGTCCCTGGGGCC 417

Db 79 GlyAsnGlnIleSerLeuLeuArgAlaPhePheTyrValAlaAlaGlnLeuValGlyAla 98

QY 418 ATCATTGGAGCAGGAATCTCTATCTGGTCACACCTCCAGCTGGTGGAGGCCCTGGGA 477

Db 99 IleAlaGlyAlaGlyIleLeuTyrGlyValAlaProLeuAsnAlaAargGlyAsnLeuAla 118

QY 478 GTCCACCATGTTTCATGGAATCTTACCGCTGGTGCATGGTCTCTCTGGTTGAGTTGATATC 537

Db 119 ValAsnAlaLeuAsnAsnAsnThrThrGlnGlyGlnAlaMetValValGluLeuIleLeu 138

QY 538 ACATTTCAATGCTGTTTACTATCTTTGCCAGCTGTGATTCCAAACGGAGTGCATCTCACT 597

Db 139 ThrPheGlnLeuAlaLeuCysIlePheAlaSerThrAspSerArgThrSerProVal 158

QY 598 GGCCTCAATAGCTTTAGCAATTTGATTTCTGTTGGCAATTTGGACATTTATTTGCAATCAAT 657

Db 159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178







RESULT 12  
 ID ADL70563 standard; protein; 265 AA.  
 XX AC ADL70563;  
 XX DT 20-MAY-2004 (first entry)  
 XX DE Cervical cancer marker OV3A, aquaporin 5 variant 1.  
 XX KW Human; cervical cancer; marker; diagnosis; gene therapy; aquaporin 5;  
 XX KW vaccine; gene therapy.  
 XX OS Homo sapiens.  
 XX PN WO2004018999-A2.  
 XX PD 04-MAR-2004.  
 XX PF 20-AUG-2003; 2003WO-US026184.  
 XX PR 20-AUG-2002; 2002US-0404770P.  
 XX PA (MILL-) MILLENIUM PHARM INC.  
 XX PI Monahan JE, Zhao X, Chen Y, Glatt K, Kamatkar S;  
 XX DR N-PSDB; ADL70562.  
 XX PT Assessing, characterizing, monitoring, preventing and treating cervical  
 PT cancer including pre-malignant conditions, e.g. dysplasia comprises  
 PT comparing the level of expression of a marker in a patient sample and in  
 PT a control sample.  
 XX PS Disclosure; SEQ ID NO 6; 167pp; English.  
 XX CC The present sequence is the protein sequence of cervical cancer marker  
 CC OV3A, characterised as human aquaporin 5 variant 1. A higher than normal  
 CC expression level of this marker correlates with the presence of cervical  
 CC cancer, including pre-malignant conditions such as dysplasia, in a  
 CC patient. The marker was identified by transcription profiling experiments  
 CC and by in situ hybridisation. The invention relates to newly discovered  
 CC cancer marker ADL70558-ADL70601 polynucleotides and polypeptides  
 CC associated with the cancerous state of cervical cells. It provides  
 CC methods, reagents and kits for diagnosing, staging, prognosing,  
 CC monitoring and treating cervical cancer (including carcinoma in situ,  
 CC invasive carcinoma and metastatic carcinoma) and pre-malignant conditions  
 CC (such as dysplasia, including CIN or SIL). A claimed method of treating  
 CC cervical cancer uses an antisense oligonucleotide that is complementary  
 CC to one of the cancer marker polynucleotides. Claimed methods of  
 CC inhibiting cervical cancer comprise administering a composition that  
 CC induces a lower expression of a marker or inhibiting expression of a  
 CC marker gene.  
 XX SQ Sequence 265 AA;  
 Alignment Scores:  
 Pred. No.: 4,27e-52 Length: 265  
 Score: 617.00 Matches: 132  
 Percent Similarity: 64.64% Conservative: 38  
 Best Local Similarity: 50.13% Mismatches: 81  
 Query Match: 28.29% Indels: 12  
 DB: 8 Gaps: 4  
 US-10-723-180-1 (1-1152) x ADL70563 (1-265)  
 QY 118 AAAGGGGTCTGCACTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTT 177  
 Db 3 LysGluValCysSerValAlaPheLeuLysAlaValPheAlaGluPheLeuAlaThrLeu 22  
 QY 178 ATTTTGTCTCTCAGCCTGGATCCACCATCACTGGGGTGAACAGAAAGCCTTTA 237

Db 23 IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp-----ProSerAlaLeu 39  
 QY 238 CCGGTCGACATGGTTCTCATCTCCCTTTGCTTTGGACTCAGCATTCGAACCATGGTCGAG 297  
 Db 40 ProThr---IleLeuGlnIleAlaLeuAlaPheGlyLeuAlaIleGlyThrLeuAlaGln 58  
 QY 298 TGCTTTGGCCATATCAGCGGTGGCCACATCAACCTCGAGTGAAGTGTGGCCATGGTGTGC 357  
 Db 59 AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuVal 78  
 QY 358 ACCAGGAAGATCAGCATCGCCAGCTCTGTCTTCTACATCGCAGCCAGTGTGGGGGCC 417  
 Db 79 GlyAsnGlnIleSerLeuLeuArgAlaPhePheYrValAlaAlaGlnLeuValGlyAla 98  
 QY 418 ATCATTTGAGCAGGAATCTCTATCTGTGTCCACCTCCCATGCTGGTGGGAGGCTGGGA 477  
 Db 99 IleAlaGlyAlaGlyIleLeuTyrGlyValAlaProLeuAsnAlaArgGlyAsnLeuAla 118  
 QY 478 GTCACCATGGTTCATGGAAATCTTACCGCTGGTTCATGCTCTCTGCTGTTGATGATATC 537  
 Db 119 ValAsnAlaLeuAsnAsnThrThrGlnGlyGlnAlaMetValValGluLeuIleLeu 138  
 QY 538 ACATTTCAATTGGTGTCTTACTCTCTTCCAGCTGTGATTCCTCAACCGGACTGATGTCAC 597  
 Db 139 ThrPheGlnLeuAlaLeuCysIlePheAlaSerThrAspSerArgArgThrSerProVal 158  
 QY 598 GGCTCAATAGCTTTAGCAATTTGGATTTTCTGTGCAATTTGGACATTTATTTGCAATCAAT 657  
 Db 159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178  
 QY 658 TATACTGTGTCACGATGAATCCCGCGATCTCTTTGGACCTGCGAGTTATCATGTTATGGAAT 717  
 Db 179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValValMetAsnArg 198  
 QY 718 TGGGAA---AACCATTGGATATTTGGTGGGCCCATCATGAGGAGCTGCTCTCGCTGGT 774  
 Db 199 PheSerProAlaHisTrpValPheTrpValGlyProIleValGlyAlaValLeuAla 218  
 QY 775 GGCCTTTATGATGATGTTCTCTGTCCAGATGTTGAAATTCAAACGTCGTTTTAAAGAGCC 834  
 Db 219 IleLeuTyrPheTyrLeuLeuPheProAsn-----SerLeuSer 231  
 QY 835 TTCAGCAAGCTGCCAGCAACAAAGGAAGCTACATGGAGGTGGAGGACACACAGGAGT 894  
 Db 232 LeuSerGluArgValAlaIleLeuLysGlyThrTyrGluProAspGluAspTrpGluGlu 251  
 QY 895 CAGGTAGAG 903  
 Db 252 GlnArgGlu 254  
 RESULT 13  
 AAW55787  
 ID AAW55787 standard; protein; 265 AA.  
 XX AC AAW55787;  
 XX DT 14-JUL-1998 (first entry)  
 XX DE Rat aquaporin-5.  
 XX KW Human; aquaporin-1; AQP-1; water channel protein; regulation;  
 KW osmotic change; erythrocyte; dryness; blindness; hydration; asthma;  
 KW secretion; aquaporin-5; AQP-5.  
 XX OS Rattus sp.  
 XX PN US5741671-A.  
 XX PD 21-APR-1998.  
 XX PF 06-JUN-1995; 95US-00468763.



Alignment Scores:		8.5e-52	Length:	265
Pred. No.:		614.00	Matches:	133
Score:		56.83%	Conservative:	46
Best Local Similarity:		42.22%	Mismatches:	82
Query Match:		28.15%	Indels:	54
DB:		2	Gaps:	5
US-10-723-180-1 (1-1152) x AAW94319 (1-265)				
QY	118	AAAGGGCTCGACCTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTTCTGCCCATGCTT	177	
DB	3	LysGluValCysSerLeuAlaPhePheLysAlaValPheAlaGluPheLeuAlaThrLeu	22	
QY	178	ATTTTGTCTCTCCTCAGCTGGGATCCACATCAACTGGGGTGGACAGAAAAGCTTTA	237	
DB	23	IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp-----ProSerAlaLeu	39	
QY	238	CGGGTCGACATGGTTCATCTCCCTTGTGGACTCAGCATTCGACCATGGTGGCAG	297	
DB	40	ProThr---IleLeuGlnIleSerIleAlaPheGlyLeuAlaIleGlyThrLeuAlaGln	58	
QY	298	TGCTTTGGCCATCATCGCGTGGCCACATCAACCTCGAGTCAGCTGTGGCCATGGTGC	357	
DB	59	AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuLeu	78	
QY	358	ACAGGAAGATCAGATCGCCCAAGTCTCTTCTCATCGCAGCCAGTCCCTGGGGGCC	417	
DB	79	GlyAsnGlnIleSerLeuLeuArgAlaValPheTyrValAlaAlaGlnLeuValGlyAla	98	
QY	418	ATCATTTGGAGGAGGATCTCTATCTGGTCACACCTCCAGTGTGGTGGAGGCTGGGA	477	
DB	99	IleAlaGlyAlaGlyIleLeuTyrTrpLeuAlaProLeuAsnAlaArgGlyAsnLeuAla	118	
QY	478	GTCCACCATGGTTCATGGAAATCTTACCGTGGTCTCATGGTCTCTCTGGTGTGATAATC	537	
DB	119	ValAsnAlaLeuAsnAsnThrThrProGlyLysAlaMetValValGluLeuLeuLeu	138	
QY	538	ACATTTCAATGGTGTACTACTTTTCCAGCTGTGATTCAAACGGACTGATCTACT	597	
DB	139	ThrPheGlnLeuAlaLeuCysIlePheSerSerThrAspSerArgArgThrSerProVal	158	
QY	598	GGCTCAATAGCTTACGATTCGATTTCTGTGCAATGGACATTTATTTCAATCAAT	657	
DB	159	GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr	178	
QY	658	TATATCTGGTCCGACATGAATCCCGCCGATCCTTTGGACCTGGCAGTTATCATGGGAAT	717	
DB	179	PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValValMetAsnArg	198	
QY	718	TGGGAA---AACATTGGANATATTTGGTTGGGCCCATCATAGGAGCTGTCTCGCTGGT	774	
DB	199	PheSerProSerHisTrpValPheTrpValGlyProIleValGlyAlaMetLeuAlaAla	218	
QY	775	GGCCTTTATGATGTCTCTGTCCAGATGTGTGAATTCAAACGTCGTTTAAAGAGCC	834	
DB	219	IleLeuTyrPheTyrLeuLeuPheProSer-----SerLeuSer	231	
QY	835	TTTCAGAAAGCTGCCAGCAAAAGGAAGCTACATGGAGGTGGAGGACCAACAGGAGT	894	
DB	232	LeuHisAspArgValAlaValValLysGlyThrTyrGluProGluGluAsp-----	248	
QY	895	CAGGTAGACGGATGACCTGATTTCTAAACCTGGAGTGGTGCATGTGATTGACGTTGAC	954	
DB	248	-----	248	
QY	955	CGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1014	
DB	249	-----	249	
QY	1015	GAAGATCGCACTGAAGCAGACAGACACTCTTAGAAGTGTCTCTCA	1059	

Db 250 GluAspHisArgGluGluArgLysThrIleGluLeuThrAla 264

RESULT 15

ADD47948

ID ADD47948 standard; protein; 265 AA.

XX AC ADD47948;

XX 02-DEC-2004 (revised)

DT 29-JAN-2004 (first entry)

XX Rat Protein AAA66221, SEQ ID NO 13644.

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

OS Unidentified.

XX WO2003016475-A2.

PD 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GHEO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENBANK; AAA66221.

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX Example 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a rat protein (described in Table 3 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 265 AA;

Alignment Scores:



This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 8, 2005, 11:56:43 ; Search time 41 Seconds  
(without alignments)  
4194.913 Million cell updates/sec

Title: US-10-723-180-1  
Perfect score: 2181  
Sequence: 1 9999caggcaatgagagctg.....attactcagctctaaacaata 1152

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p.model -DEV=xlp  
-O=/cn2 1/USPTO spool\_p/US10723180/runat 08092005 123257 29261/app query.fasta\_1.1351  
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10723180 @CN 1.1 46 @runat 08092005 123257 29261 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DRV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:\*  
1: /cn2 6/ptodata/1/iaa/5A COMB pep:\*  
2: /cn2 6/ptodata/1/iaa/5B COMB pep:\*  
3: /cn2 6/ptodata/1/iaa/6A COMB pep:\*  
4: /cn2 6/ptodata/1/iaa/6B COMB pep:\*  
5: /cn2 6/ptodata/1/iaa/PCTUS COMB pep:\*  
6: /cn2 6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1755	80.5	336	4	US-09-949-016-10148
2	1755	80.5	336	4	US-09-949-016-10149
3	624.5	28.6	437	4	US-09-949-016-8094
4	614	28.2	265	1	US-08-468-763-19
5	614	28.2	265	2	US-08-393-996A-19
6	602	28.1	293	4	US-09-949-016-8027
7	603	27.6	118	4	US-09-513-999C-6089
8	594	27.2	271	1	US-08-447-554-4
9	594	27.2	271	1	US-08-448-160-4
10	561	25.7	269	1	US-08-468-763-17
11	561	25.7	269	2	US-08-393-996A-17
12	561	25.7	269	4	US-09-949-016-5934

13	561	25.7	281	4	US-09-949-016-8317	Sequence 8317, Ap
14	561	25.7	281	4	US-09-949-016-10471	Sequence 10471, A
15	560	25.7	269	1	US-08-447-554-5	Sequence 5, Appli
16	560	25.7	269	1	US-08-448-160-5	Sequence 5, Appli
17	426.5	19.6	284	3	US-09-372-448A-4	Sequence 4, Appli
18	419	19.2	288	3	US-09-372-422A-12	Sequence 12, Appli
19	418	19.2	292	3	US-09-372-422A-4	Sequence 4, Appli
20	417	19.1	288	3	US-09-372-448A-2	Sequence 2, Appli
21	417	19.1	289	3	US-09-372-422A-2	Sequence 2, Appli
22	415	19.0	289	3	US-09-372-422A-14	Sequence 14, Appli
23	414	19.0	288	3	US-09-372-422A-18	Sequence 18, Appli
24	410	18.8	296	3	US-09-372-422A-20	Sequence 20, Appli
25	406.5	18.6	292	3	US-09-372-422A-10	Sequence 10, Appli
26	398.5	18.3	254	3	US-09-372-422A-34	Sequence 34, Appli
27	395	18.1	249	3	US-09-372-422A-22	Sequence 22, Appli
28	390.5	17.9	288	3	US-09-372-422A-16	Sequence 16, Appli
29	388	17.8	250	3	US-08-654-025-2	Sequence 2, Appli
30	387	17.7	250	1	US-08-234-939-2	Sequence 2, Appli
31	387	17.7	250	1	US-08-558-865-2	Sequence 2, Appli
32	387	17.7	250	3	US-08-654-025-7	Sequence 7, Appli
33	384.5	17.6	272	3	US-09-372-422A-26	Sequence 26, Appli
34	377.5	17.3	257	3	US-09-372-422A-28	Sequence 28, Appli
35	375	17.2	262	3	US-09-372-422A-32	Sequence 32, Appli
36	364	16.7	249	3	US-09-372-448A-6	Sequence 6, Appli
37	358	16.4	281	3	US-09-053-702-2	Sequence 2, Appli
38	353.5	16.2	249	3	US-09-372-422A-30	Sequence 30, Appli
39	352.5	16.2	250	3	US-09-372-422A-24	Sequence 24, Appli
40	334	15.3	247	3	US-09-372-422A-48	Sequence 48, Appli
41	332.5	15.2	294	3	US-09-372-422A-40	Sequence 40, Appli
42	320.5	14.7	295	3	US-09-372-422A-38	Sequence 38, Appli
43	311.5	14.3	263	4	US-09-610-906-12	Sequence 12, Appli
44	306.5	14.1	262	4	US-09-976-594-347	Sequence 347, App
45	306.5	14.1	443	4	US-09-610-906-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-949-016-10148  
; Sequence 10148, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE OF INVENTION: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10148  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-10148

Alignment Scores:	2.89e-180	Length:	336
Pred. No.:	1755.00	Matches:	336
Score:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.47%	Indels:	0
DB:	4	Gaps:	0

US-10-723-180-1 (1-1152) x US-09-949-016-10148 (1-336)

QY 1 GGGCAGGCAATGAGAGCTGCACTCTGGCTGGGAAGGCATGAGTGACAGCCACAGCA 60

```
Db 1 GlyAlaGlyAsnGluSerCysThrLeuAlaGlyGluGlyMetSerAspArgProThrAla 20
QY 61 AGCGGTGGGTAAAGTGTGACCTTTGTGTACAGAGAGAAACATCATGTGTGCTTCAAA 120
Db 21 ArgAgtTrpGlyLysCysGlyProLeuCysThrArgGluAsnIleMetValAlaPheLys 40
QY 121 GGGGTCTGACACTCAAGCTTCTGGAAGCAGNACAGCGGAATTTCTGGCCATGCTTATT 180
Db 41 GlyValTrpThrGlnAlaPheTrpLysAlaValThrAlaGluPheLeuAlaMetLeuIle 60
QY 181 TTTGTCTCTCCTCAGCCTCGGATCCACCATCACTGGGTGGAAACAGAAAGCCTTTACCG 240
Db 61 PheValLeuLeuSerLeuGlySerThrIleAsnTrpGlyGlyThrGluLysProLeuPro 80
QY 241 GTCGACATGTTCTCATCTCCCTTTGCTTTGTGACTCAGCATTTGCAACACCATGTCAGTGC 300
Db 81 ValAspMetValLeuIleSerLeuGlyPheGlyLeuSerIleAlaThrMetValGlnCys 100
QY 301 TTTGGCCATATCAGCGGTGGCCACATCAACCTGCAGTCACTGTGGCCATGTTGTCACC 360
Db 101 PheGlyHisIleSerGlyGlyHisIleAsnProAlaValThrValAlaMetValCysThr 120
QY 361 AGGAAGATCAGCATCGCCAAAGTCTGTCTCTACATCGCAGCCAGTGCCTGGGGCCATC 420
Db 121 ArgLysIleSerIleAlaLysSerValPheThrIleAlaAlaGlnCysLeuGlyAlaIle 140
QY 421 ATTGAGCAGGAATCCTCTATCTGTCTACACCTCCCAAGTGTGGTGGAGCGCTGGAGTGC 480
Db 141 IleGlyAlaGlyIleLeuTrpValThrProSerValValGlyGlyLeuGlyVal 160
QY 481 ACCATGTTTCATGGAATCTTACCGCTGTCTGCTGTCTCTGCTGTGAGTTGATATCA 540
Db 161 ThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeuLeuValGluLeuIleIleThr 180
QY 541 TTTCAATTGCTTTACTATCTTTCAGCTGTGATTCCAAACGAGTCACTGTCTCACTGTC 600
Db 181 PheGlnLeuValPheThrIlePheAlaSerCysAspSerLysArgThrAspValThrGly 200
QY 601 TCAATAGCTTTAGCAATTTGATTTTCTGTGCAATTTGACATTTATTTGCAATCAATTAT 660
Db 201 SerIleAlaLeuAlaIleGlyPheSerValAlaIleGlyHisLeuPheAlaIleAsnTrp 220
QY 661 ACTGTCAGCATGATGATCCCGCATCTTTGACCTGTCAGTCACTATCAATGGAAATGG 720
Db 221 ThrGlyAlaSerMetAsnProAlaArgSerPheGlyProAlaValIleMetGlyAsnTrp 240
QY 721 GAAACCATTTGATATTTGGGTGGCCCATCATAGGAGCTGTCTCTGCTGTCGCTT 780
Db 241 GluAsnHisTrpIleTrpValGlyProIleIleGlyAlaValLeuAlaGlyGlyLeu 260
QY 781 TATGATATGCTCTCTGTCAGATGTTGAATTCAAACGTCGTTTAAAGAACCTTCAGC 840
Db 261 TyrGluTrpValPheCysProAspValGluPheLysArgPheLysGluAlaPheSer 280
QY 841 AAAGCTGCCAGCAACAAAGAGAGCTACATGAGGTGGAGGACACAGGAGTCAAGTA 900
Db 281 LysAlaAlaGlnGlnThrLysGlySerTyrMetGluValGluAspAsnArgSerGlnVal 300
QY 901 GAGACGGATGACTGATTTCTAAACCTGGAGTGGTGCATGTGATTCACGTTTACCGGGGA 960
Db 301 GluThrAspAspLeuIleLysProGlyValValHisValIleAspValAspArgGly 320
QY 961 GAGGAGAAAGGGGAAAGCAATCTGGAGAGGATTTGCTTTCAGTA 1008
Db 321 GluGluLysGlyLysAspGlnSerGlyGluValLeuSerSerVal 336
```

## RESULT 2

```
US-09-949-016-10149
; Sequence 10149, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10149
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10149
```

```
Alignment Scores:
Pred. No.: 2,89e-180 Length: 336
Score: 1755.00 Matches: 336
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.47% Indels: 0
DB: 4 Gaps: 0
```

```
US-10-723-180-1 (1-1152) x US-09-949-016-10149 (1-336)
```

```
QY 1 GGGCAGGCAATAGAGCTGCACCTGTGGCTGGGAAGGCATGATGACAGACCCACGACGA 60
Db 1 GlyAlaGlyAsnGluSerCysThrLeuAlaGlyGluGlyMetSerAspArgProThrAla 20
QY 61 AGCGGTGGGTAAAGTGTGACCTTTGTGTACAGAGAGAAACATCATGTGTGCTTCAAA 120
Db 21 ArgAgtTrpGlyLysCysGlyProLeuCysThrArgGluAsnIleMetValAlaPheLys 40
QY 121 GGGGTCTGACACTCAAGCTTCTGGAAGCAGTCAACCGGTGGAAATTTCTGGCCATGCTTATT 180
Db 41 GlyValTrpThrGlnAlaPheTrpLysAlaValThrAlaGluPheLeuAlaMetLeuIle 60
QY 181 TTTGTCTCTCCTCAGCCTGGATCCACCATCACTGGGTGGAAACAGAAAGCCTTTACCG 240
Db 61 PheValLeuLeuSerLeuGlySerThrIleAsnTrpGlyGlyThrGluLysProLeuPro 80
QY 241 GTCGACATGTTCTCATCTCCCTTTGCTTTGTGACTCAGCATTTGCAACACCATGTCAGTGC 300
Db 81 ValAspMetValLeuIleSerLeuGlyPheGlyLeuSerIleAlaThrMetValGlnCys 100
QY 301 TTTGGCCATATCAGCGGTGGCCACATCAACCTGCAGTCACTGTGGCCATGTTGTCACC 360
Db 101 PheGlyHisIleSerGlyGlyHisIleAsnProAlaValThrValAlaMetValCysThr 120
QY 361 AGGAAGATCAGCATCGCCAAAGTCTGTCTCTACATCGCAGCCAGTGCCTGGGGCCATC 420
Db 121 ArgLysIleSerIleAlaLysSerValPheThrIleAlaAlaGlnCysLeuGlyAlaIle 140
QY 421 ATTGAGCAGGAATCCTCTATCTGTGTCACACCTCCCAAGTGTGGTGGAGCGCTGGAGTGC 480
Db 141 IleGlyAlaGlyIleLeuTrpValThrProSerValValGlyGlyLeuGlyVal 160
QY 481 ACCATGTTTCATGGAATCTTACCGCTGTCTGCTGTCTCTGCTGTGAGTTGATATCA 540
Db 161 ThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeuLeuValGluLeuIleIleThr 180
QY 541 TTTCAATTGCTTTACTATCTTTCAGCTGTGATTCCAAACGAGTCACTGTCTCACTGTC 600
Db 181 PheGlnLeuValPheThrIlePheAlaSerCysAspSerLysArgThrAspValThrGly 200
QY 601 TCAATAGCTTTAGCAATTTGATTTTCTGTGCAATTTGACATTTATTTGCAATCAATTAT 660
Db 201 SerIleAlaLeuAlaIleGlyPheSerValAlaIleGlyHisLeuPheAlaIleAsnTrp 220
```

```
QY 661 ACTGGTCCAGCATGAATCCCGCCGATCCTTTGGACCTGACATTCATATGCGAAATTGG 720
Db 221 ThrGlyAlaSerMetAsnProAlaArgSerPheGlyProAlaValIleMetGlyAsnTyr 240
QY 721 GAAACCATTTGGATATTTGGTGGCCCATCATAGGAGCTGCTCCCTGGTGGCCCTT 780
Db 241 GluAsnHisTrpIleTyrTrpValGlyProIleIleGlyAlaValIleAlaGlyGlyLeu 260
QY 781 TATGAGTATGTCTTCTGCTCCAGATGTTGAATTCAAACGCTGTTTAAAGAAGCCTTCAGC 840
Db 261 TyrGluTyrValPheCysProAspValGluPheLysArgPheLysGluAlaPheSer 280
QY 841 AAGCTGCCAGCAACAAGAGCACTCATCGAGGTGGAGCAACAAGGAGTCAAGTA 900
Db 281 LysAlaAlaGlnInThrLysGlySerTyrMetGluValGluAspAsnArgSerGlnVal 300
QY 901 GAGACGGATGACCTGATTTCTAAACCTGGAGTGGTGCATGTGATTCAGCTTGACCGGGA 960
Db 301 GluThrAspAspLeuIleLeuLysProGlyValValHisValIleAspValAspArgGly 320
QY 961 GAGGAGAAGAGGGAAGACCAATCTGGAGAGTATTGTCTTCAGTA 1008
Db 321 GluGluLysLysGlyLysAspGlnSerGlyGluValLeuSerSerVal 336

RESULT 3
US-09-949-016-8094
; Sequence 8094, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8094
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8094

Alignment Scores:
Pred. No.: 2,85e-58 Length: 437
Score: 624.50 Matches: 140
Percent Similarity: 62.15% Conservative: 39
Best Local Similarity: 48.61% Mismatches: 88
Query Match: 28.63% Indels: 21
DB: 4 Gaps: 7

US-10-723-180-1 (1-1152) x US-09-949-016-8094 (1-437)
QY 52 CCACA---GCAAGCGGTGGGTGAAGTGT-----GGACCTTTGTGTACAGAGAAC 102
Db 156 ProThrGlyAlaProArgGlySerCysArgAlaProAlaAlaThrMetLys--- 174
QY 103 ATCATGGTGGCTTCAAGGGGTCTGGACTCAAGCTTCTGGAAGAGCTACAGGGAA 162
Db 175 -----LysGluValCysSerValAlaPheLeuLysAlaValPheAlaGlu 189
QY 163 TTTCTGGCCATGCTTATTTTGTCTCTCCCTCAGCTGGGATCCACCATCAACTGAGGTGA 222
Db 190 PheLeuAlaThrIleIlePheValPhePheGlyLeuGlySerAlaLeuLysTrp----- 207
QY 223 ACAGAAAAGCCTTTACCGGTGCACATGGTTCTCATCTCCCTTTGTGACTTCAGCAT 282
```

```
Db 208 ---ProSerAlaLeuProThr---IleLeuGlnIleAlaLeuAlaPheGlyLeuAlaIle 225
QY 283 GCAACCATGGTGCAGTCTTTGGCCATATACAGGGTGGCCACATCAACACCTGCAGTGACT 342
Db 226 GlyThrLeuAlaGlnAlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThr 245
QY 343 GTGGCCATGGTGTGCACCAGGAAGATCAGCATCGCAAGTCTGTCTTCTACATCGCAGCC 402
Db 246 LeuAlaLeuLeuValGlyAsnGlnIleSerLeuLeuArgAlaPhePheTyrValAlaAla 265
QY 403 CAGTGCCTGGGGCCCATCATTCGAGCAGGAATCTCTATCTGTGTACACCTCCAGTGTG 462
Db 266 GlnLeuValGlyAlaIleAlaGlyIleLeuTyrGlyValAlaProLeuAsnAla 285
QY 463 GTGGGAGGCTCGGGAGTCCACCATGTTTCATGAAATCTTTACCGCTGGTCTGCTCTCC 522
Db 286 ArgGlyAsnLeuAlaValAsnAlaLeuAsnAsnThrThrGlnGlyGlnAlaMetVal 305
QY 523 GTTGAAGTTGAATATCAATTTCAATTTGGTGTGTATCTATCTTTGCCAGCTGTGATTCAAA 582
Db 306 ValGluLeuIleLeuThrPheGlnLeuAlaLeuCysIlePheAlaSerThrAspSerArg 325
QY 583 CGGACTGATGTCAGTGGCTCAATAGCTTTAGCAATTTGGAATTTCTGTTCGAATTTGCAAT 642
Db 326 ArgThrSerProValGlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHis 345
QY 643 TTTATTTGCAATCAATATATCTAGTCCGAGCATGAATCCCGCCGATCCTTTGAGCTGCA 702
Db 346 LeuValGlyIleTyrPheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAla 365
QY 703 GTTATCTATCGGAAATTTGGAA---AACCATTTGGATATATTGGTTGGGCCCATCATAGGA 759
Db 366 ValValMetAsnArgPheSerProAlaHisTrpValPheTrpValGlyProIleValGly 385
QY 760 GCTGTCTCGCTGGTGGCTTTATGAGTATGCTCTCTGTCAGATGTTGGAATTTCAAACT 819
Db 386 AlaValLeuAlaAlaIleLeuTyrPheTyrLeuLeuPheProAsn----- 400
QY 820 CGTTTTAAAGAGCCTTCAGCAAGCTGCCCAAGCTGCCCAAGCAACAAAGAGAGCTACATGGAGGTG 879
Db 401 -----SerLeuSerLeuSerGluArgValAlaIleLysGlyThrTyrGluProAsp 418

RESULT 4
US-08-468-763-19
; Sequence 19, Application US/08468763
; Patent No. 5741671
; GENERAL INFORMATION:
; APPLICANT: Agre, Peter C.
; TITLE OF INVENTION: Isolation, Cloning and Expression of
; TITLE OF INVENTION: Transmembrane Water Channel Proteins
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: D.C.
; COUNTRY: US
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,763
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,996
```

```
/ FILING DATE: 24-FEB-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Posorske, Laurence H.
/ REGISTRATION NUMBER: 34,698
/ REFERENCE/DOCKET NUMBER: 1107.48633
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202 508-9100
/ TELEFAX: 202 508-9299
/ INFORMATION FOR SEQ ID NO: 19:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 265 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-468-763-19

Alignment Scores:
Pred. No.: 2,93e-57 Length: 265
Score: 614.00 Matches: 133
Percent Similarity: 56.83% Conservatives: 46
Best Local Similarity: 42.22% Mismatches: 82
Query Match: 28.15% Indels: 54
DB: 1 Gaps: 5

US-10-723-180-1 (1-1152) x US-08-468-763-19 (1-265)

QY 118 AAAGGGTCTGCACTCAAGCTTTCTGGAAGAGCTACATGGAGGTGGAGGCAACAGGAGT 894
Db 118 AAAGGGTCTGCACTCAAGCTTTCTGGAAGAGCTACATGGAGGTGGAGGCAACAGGAGT 894
QY 3 LysGluValCysSerLeuAlaPhePheLysAlaValPheAlaGluPheLeuAlaThrLeu 22
Db 3 LysGluValCysSerLeuAlaPhePheLysAlaValPheAlaGluPheLeuAlaThrLeu 22
QY 178 ATTTTGTCTCTCAGCTGGAGTCCACCATCACTGGGTGGACAGAAAAGCCTTTA 237
Db 178 ATTTTGTCTCTCAGCTGGAGTCCACCATCACTGGGTGGACAGAAAAGCCTTTA 237
QY 23 IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp-----ProSerAlaLeu 39
Db 23 IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp-----ProSerAlaLeu 39
QY 238 CGGTCGACATGGTTCTCATCTCCCTTTCTGTTGACTCAGCATTCGAACCATGGTGCAG 297
Db 238 CGGTCGACATGGTTCTCATCTCCCTTTCTGTTGACTCAGCATTCGAACCATGGTGCAG 297
QY 40 ProThr---IleLeuGlnIleSerIleAlaPheGlyLeuAlaIleGlyThrLeuAlaGln 58
Db 40 ProThr---IleLeuGlnIleSerIleAlaPheGlyLeuAlaIleGlyThrLeuAlaGln 58
QY 298 TCGTTTGGCCATATCAGCGGTGGCCATCAACCTTGCAGTCACTGTGGCCATGGTGC 357
Db 298 TCGTTTGGCCATATCAGCGGTGGCCATCAACCTTGCAGTCACTGTGGCCATGGTGC 357
QY 59 AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuLeu 78
Db 59 AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuLeu 78
QY 358 ACCGAGATCAGCATGCCAAGTCTGCTTCTACATCGCAGCCAGTCCCTGGGGGCC 417
Db 358 ACCGAGATCAGCATGCCAAGTCTGCTTCTACATCGCAGCCAGTCCCTGGGGGCC 417
QY 79 GlyAsnGlnIleSerLeuAlaValPheTrpValAlaAlaGlnLeuValGlyAla 98
Db 79 GlyAsnGlnIleSerLeuAlaValPheTrpValAlaAlaGlnLeuValGlyAla 98
QY 418 ATCATTGGAGAGGAATCTCTATCTGTGTGTCACACCTCCAGCTGTGTGGAGGCTGGGA 477
Db 418 ATCATTGGAGAGGAATCTCTATCTGTGTGTCACACCTCCAGCTGTGTGGAGGCTGGGA 477
QY 99 IleAlaGlyAlaGlyIleLeuTrpLeuAlaProLeuAsnAlaArgGlyAsnLeuAla 118
Db 99 IleAlaGlyAlaGlyIleLeuTrpLeuAlaProLeuAsnAlaArgGlyAsnLeuAla 118
QY 478 GTCACCATGGTTCATGGAAATCTTACCGCTGGTCAATGCTCTCTGTTGAGTTGATATC 537
Db 478 GTCACCATGGTTCATGGAAATCTTACCGCTGGTCAATGCTCTCTGTTGAGTTGATATC 537
QY 119 ValAsnAlaLeuAsnAsnAsnThrProGlyLysAlaMetValValGluLeuLeu 138
Db 119 ValAsnAlaLeuAsnAsnAsnThrProGlyLysAlaMetValValGluLeuLeu 138
QY 538 ACATTTCAATTTGTTTACTATCTTCCAGCTGTGATTCACAAAGGAGCTGATCTACT 597
Db 538 ACATTTCAATTTGTTTACTATCTTCCAGCTGTGATTCACAAAGGAGCTGATCTACT 597
QY 139 ThrPheGlnLeuAlaLeuCysIlePheSerSerThrAspSerArgArgThrSerProVal 158
Db 139 ThrPheGlnLeuAlaLeuCysIlePheSerSerThrAspSerArgArgThrSerProVal 158
QY 598 GGCTCAATAGCTTTAGCAATTTGATTTCTGTTGCAATTTGACATTTATTTGCAATCAAT 657
Db 598 GGCTCAATAGCTTTAGCAATTTGATTTCTGTTGCAATTTGACATTTATTTGCAATCAAT 657
QY 159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178
Db 159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178
QY 658 TATACTGGTCCAGCATGAAATCCCGCCGATCCTTTGGACCTGCGAGTTATCATGGGAAT 717
Db 658 TATACTGGTCCAGCATGAAATCCCGCCGATCCTTTGGACCTGCGAGTTATCATGGGAAT 717
QY 179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValValMetAsnArg 198
Db 179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValValMetAsnArg 198
QY 718 TGGGAA---AACCTTGGATATATGGGTTGGCCCATCATAGGAGCTGCTCTCGCTGT 774
Db 718 TGGGAA---AACCTTGGATATATGGGTTGGCCCATCATAGGAGCTGCTCTCGCTGT 774
QY 199 PheSerProSerHisTrpValPheTrpValGlyProIleValGlyAlaMetLeuAla 218
Db 199 PheSerProSerHisTrpValPheTrpValGlyProIleValGlyAlaMetLeuAla 218
QY 775 GGCCTTTATGATGTCTCTGTCCAGATGTTGAATTCAAACGCTGTTTAAAGAGCC 834
Db 775 GGCCTTTATGATGTCTCTGTCCAGATGTTGAATTCAAACGCTGTTTAAAGAGCC 834
QY 219 IleLeuTrpPheTrpLeuLeuPheProSer-----SerLeuSer 231
Db 219 IleLeuTrpPheTrpLeuLeuPheProSer-----SerLeuSer 231
```

```
QY 835 TTCAGCAAGCTGCCAGCAAAACAAAGAGAGCTACATGGAGGTGGAGGCAACAGGAGT 894
Db 232 LeuHisAspArgValAlaValValLysGlyThrTyrGluProGluGluAsp----- 248
QY 895 CAGGTAGAGACGGATGACCTGATTCTTAAACCTCGAGTGGTGCATGTGATTCACCTTGCAC 954
Db 248 ----- 248
QY 955 CGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1014
Db 249 -----Trp 249
QY 1015 GAAGATCCACTGAAAGCAGACAGACACTCCTTAGAAGTCTCTCA 1059
Db 250 GluAspHisArgGluGluArgLysLysThrIleGluLeuThrAla 264
```

```
RESULT 5
US-08-393-996A-19
; Sequence 19, Application US/08393996A
; Patent No. 5858702
; GENERAL INFORMATION:
; APPLICANT: Agre, Peter C.
; TITLE OF INVENTION: Isolation, Cloning and Expression of
; TRANSMEMBRANE WATER CHANNEL PROTEINS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESS: Banner & Allegretti
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: D.C.
; COUNTRY: US
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,996A
; FILING DATE: 24-FEB-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske, Laurence H.
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 1107.48633
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 508-9100
; TELEFAX: 202 508-9299
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 265 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
/ US-08-393-996A-19
```

```
Alignment Scores:
Pred. No.: 2,93e-57 Length: 265
Score: 614.00 Matches: 133
Percent Similarity: 56.83% Conservatives: 46
Best Local Similarity: 42.22% Mismatches: 82
Query Match: 28.15% Indels: 54
DB: 2 Gaps: 5

US-10-723-180-1 (1-1152) x US-08-393-996A-19 (1-265)

QY 118 AAAGGGTCTGCACTCAAGCTTTCTGGAAGAGCTACATGGAGGTGGAGGCAACAGGAGT 177
Db 3 LysGluValCysSerLeuAlaPhePheLysAlaValPheAlaGluPheLeuAlaThrLeu 22
QY 178 ATTTTGTCTCTCAGCTGGAGTCCACCATCACTGGGTGGACAGAAAAGCCTTTA 237
Db 178 ATTTTGTCTCTCAGCTGGAGTCCACCATCACTGGGTGGACAGAAAAGCCTTTA 237
```

```
Db 23 IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp-----ProSerAlaLeu 39
QY 238 CCGGTGCGACATGTTCTCATCTCCCTTCTTGGACTCAGCATTCGAACCATGTGCGAG 297
Db 40 ProThr--IleLeuGlnIleSerIleAlaPheGlyLeuAlaIleGlyThrLeuAlaGln 58
QY 298 TCGCTTTGGCCATATCAGCGGTGGCCACATCAACCTTCGAGTGAAGTGTGCCATGTGTC 357
Db 59 AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuLeuIle 78
QY 358 ACCAGGAGATCAGCATCCCAAGTCTCTCTCATCGCAGCCGACGTCCTGGGGCC 417
Db 79 GlyAsnGlnIleSerLeuLeuAlaValPheTyrValAlaAlaGlnLeuValGlyAla 98
QY 418 ATCATTTGGAGCAGGAATCTCTATCTGTGTGCACACCTCCAGCTGTGTGGAGGCTGGGA 477
Db 99 IleAlaGlyAlaGlyIleLeuTyrTrpLeuAlaProLeuAsnAlaArgGlyAsnLeuAla 118
QY 478 GTCCACATGGTTCATGGAAATCTTACCGCTGTGTCTCATCGCAGCCGACGTCCTGGTGAATGATAATC 537
Db 119 ValAsnAlaLeuAsnAsnThrThrProGlyLysAlaMetValValGluLeuIleLeu 138
QY 538 ACATTTCAATGGTGTCTTACTATCTTGGCAGCTGTGATTCCAAACGGACTGATGTCAC 597
Db 139 ThrPheGlnLeuAlaLeuCysIlePheSerSerThrAspSerArgArgThrSerProVal 158
QY 598 GGCTCAATAGCTTTAGCAATTCGATTTCTGTGCAATTCGACATTTATTTGCAATCAAT 657
Db 159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178
QY 658 TATATCTGGTCCAGCATGAATCCCGCCGATCCTTTGGACCTGCAGTTATCATGGAAT 717
Db 179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValMetAsnArg 198
QY 718 TGGGAA--AACCATGGATATATGGGTGGGCCCATCATAGGAGCTCTCTCCTGCT 774
Db 199 PheSerProSerHisTrpValPheTrpValGlyProIleValGlyAlaMetLeuAla 218
QY 775 GGCTTTTATGATGATGCTCTCTGTCCAGATGTTGAATTCAAACGTCGTTTTAAAGAGCC 834
Db 219 IleLeuTyrPheTyrLeuLeuPheProSer-----SerLeuSer 231
QY 835 TTCAGCAAGCTGCCAGCAAAACAAAGGAAGCTATCAGAGTGGAGGACCAACAGGAGT 894
Db 232 LeuHisAspArgValAlaValValLysGlyThrTyrGluProGluGluAsp----- 248
QY 895 CAGGTAGACACGGATGACCTGATTTCTAAACCTGGAGTGGTGCATGTCATGACGTTGAC 954
Db 248 ----- 248
QY 955 CGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1014
Db 249 -----Trp 249
QY 1015 GAAGATCGCAGCTGAAGACAGACAGACAGCTCTTTAGAACTGTCCTCA 1059
Db 250 GluAspHisArgGluGluArgLysLysThrIleGluLeuThrAla 264
RESULT 6
US-09-949-016-8027
; Sequence 8027, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```

```
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8027
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8027

Alignment Scores:
Pred. No.: 5,09e-57 Length: 293
Score: 612.00 Matches: 114
Percent Similarity: 65.27% Conservative: 57
Best Local Similarity: 43.51% Mismatches: 77
Query Match: 28.06% Indels: 14
DB: 4 Gaps: 3

US-10-723-180-1 (1-1152) x US-09-949-016-8027 (1-293)
QY 124 GTCTGG-----ACTCAAGTCTTCTGGAAGCAGTCACAGCGGAATTTCTGGCCATG 174
Db 31 MetTrpGluLeuArgSerAlaSerPheTrpArgAlaIlePheAlaGluPheAlaThr 50
QY 175 CTTATTTTGTCTCTCAGCGCTGGGATCCACCATCAACTGGGGTGGAAAGAGCCT 234
Db 51 LeuPheTyrValPhePheGlyLeuGlySerSerLeuArgTrpAla-----Pro 66
QY 235 TTACCGTGCAGATGGTCTCATCTCCCTTGTCTTTGGACTCAGCATTCGAACCATGTTG 294
Db 67 GlyProLeuHisValLeuGlnValAlaMetAlaPheGlyLeuAlaLeuAlaThrLeuVal 86
QY 295 CAGTCTTTGGCCATATCAGCGGTGGCCACATCAACCTCGAGTCAGTCAGTGGCCATGTTG 354
Db 87 GlnSerValGlyHisIleSerGlyAlaHisValAsnProAlaValThrPheAlaPheLeu 106
QY 355 TGCACCAAGAGATCAGCATCGCCAAAGTCTGTCTTCTACATCGCAGCCAGTCGCTGGGG 414
Db 107 ValGlySerGlnMetSerLeuLeuArgAlaPheCysTyrMetAlaAlaGlnLeuGly 126
QY 415 GCCATCATGGAGCAGGAATCTCTATCTGTGTCACACCTCCAGTGTGTGGGAGCCCTG 474
Db 127 AlaValAlaGlyAlaValLeuTyrSerValThrProProAlaValArgGlyAsnLeu 146
QY 475 GGAGTCACCATGTTTCATGGAAATCTTACCGCTGTCTCTCTCTCTCTCTCTCTCTCTCT 534
Db 147 AlaLeuAsnThrLeuHisProAlaValSerValGlyGlnAlaThrValGluIlePhe 166
QY 535 ATCATTTCATTTGGTGTCTTCTATCTTTGCCAGCTGTGATTCCAAAGGAGCTGATGTC 594
Db 167 LeuThrLeuGlnPheValLeuCysIlePheAlaThrTyrAspGluArgAsnGlyGln 186
QY 595 ACTGGCTCAATAGCTTTAGCAATTTGATTTCTGTGCAATTTGGCAATTTATTTGCAATC 654
Db 187 LeuGlySerValAlaLeuAlaValGlyPheSerLeuAlaLeuGlyHisLeuPheGlyMet 206
QY 655 AATTATATCTGTGCCAGCATCAATCCCGCCGATCCTTTGGACCTCTTTCAGTTCATGGA 714
Db 207 TyrTyrThrGlyAlaGlyMetAsnProAlaArgSerPheAlaProAlaIleLeuThrGly 226
QY 715 AATGGGAAAAACCATTTGGATATTTGGTGTGGGCCCATCATAGGAGCTGTCTCTCGTGT 774
Db 227 AsnPheThrAsnHisTrpValTyrTrpValGlyProIleIleGlyGlyLeuGlySer 246
QY 775 GGCCTTTATAGTATGCTCTCTCTCCAGATGTTGAATTCAAACGTCGTTTAAAGAGCC 834
Db 247 LeuLeuTyrAspPheLeuLeuPheProArgLeu-----LysSer 259
QY 835 TTCACCAAGCTGCCCAACCAACAAAGGAGCTACATGGAGTGGAGGACCAACAGGAGT 894
Db 260 IleSerGluArgLeuSerValLeuLysGlyAlaLysProAspValSerAsnGlyGlnPro 279
QY 895 CAGGTA 900
```

```
Db 280 GluVal 281
:::|||||
Db 101 ThrValAlaMetValCysThrArg**IleSerIleAlaIysSerValPheTyr 118

RESULT 7
US-09-513-999C-6089
; Sequence 6089, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59 US2,REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6089
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 57
; OTHER INFORMATION: Xaa=Ile or Met
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 58
; OTHER INFORMATION: Xaa=Asp or Asn
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 75
; OTHER INFORMATION: Xaa=Leu or Arg
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 109
; OTHER INFORMATION: Xaa=Glu or Lys
US-09-513-999C-6089

Alignment Scores:
Pred. No.: 2,866-56 Length: 118
Score: 603.00 Matches: 114
Percent Similarity: 96.61% Conservative: 0
Best Local Similarity: 96.61% Mismatches: 4
Query Match: 27.65% Indels: 0
DB: 4 Gaps: 0

US-10-723-180-1 (1-1152) x US-09-513-999C-6089 (1-118)
QY 40 ATGAGTGCAGACCCACAGAGCGGTGGGTAAAGTGTGGACCTTTGTGTACGAGAG 99
Db 1 MetSerAspArgProThrAlaArgArgTrpGlyLysCysGlyProLeuCysThrArgGlu 20
QY 100 AACATCATGGTGGCTTCAAAGGGGTGGACTCAAGCTTCTGGAAGACGATCACGG 159
Db 21 AsnIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAla 40
QY 160 GAATTTCTGGCCATGCTATTATTTTGTCTCCTCAGCTGGGATCCACCATCACTGGGT 219
Db 41 GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThr*****TrpGly 60
QY 220 GGAACAGAAACCTTTACCGGTGCACATGGTTCTCATCTCCCTTTGCTTGGACTCAGC 279
Db 61 GlyThrGluLysProLeuProValAspMetValLeuIleSer**CysPheGlyLeuSer 80
QY 280 ATTGCAACCATGGTCAGTCTTTGGCCATATCAGGGTGGCCACATCAACCTCAGTG 339
Db 81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100
QY 340 ACTGTGGCCATGGTGTGCACAGGAGATCAGCATCGCCAAAGTCTGTCTTCTAC 393

Db 101 ThrValAlaMetValCysThrArg**IleSerIleAlaIysSerValPheTyr 118

RESULT 8
US-08-447-554-4
; Sequence 4, Application US/08447554
; Patent No. 5661003
; GENERAL INFORMATION:
; APPLICANT: FUSHIMI, KIYOHIDE
; APPLICANT: UCHIDA, SHINICHI
; APPLICANT: SASAKI, SEI
; APPLICANT: MARUMO, FUMIAKI
; TITLE OF INVENTION: WATER CHANNEL
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. NW, Ste. 5500
; CITY: Washington, DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,554
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,365
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 5100-0003.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-447-554-4

Alignment Scores:
Pred. No.: 4,266-55 Length: 271
Score: 594.00 Matches: 111
Percent Similarity: 66.67% Conservative: 39
Best Local Similarity: 49.33% Mismatches: 71
Query Match: 27.24% Indels: 4
DB: 1 Gaps: 1

US-10-723-180-1 (1-1152) x US-08-447-554-4 (1-271)
QY 136 GCTTCTCGAAAGCAGTCACAGCGAAATTTCTGGCCATGCTTTTGTCTCTCTCAGC 195
Db 8 AlaPheSerArgAlaValLeuAlaGluPheLeuAlaThrLeuLeuPheValPheGly 27
QY 196 CTGGATCCACATCACTGGGGTGGAGAACAAAGCCCTTACCGGTGCACATGGTTCTC 255
Db 28 LeuGlySerAlaLeuGlnTrpAlaSerSer*****ProProSerValLeuGln 43
QY 256 ATCTCCCTTCTGTTGGACTCAGACTGCAACCATGGTGCAGTGTCTTTGGCCATATCAGC 315
Db 44 IleAlaValAlaPheGlyLeuGlyIleGlyLeuValGlnAlaLeuGlyHisValSer 63
QY 316 GGTGCCACATCAACCTCGAGTCACTGTGGCCATGGTGTGCACAGGAGATCAGCATC 375
Db 64 GlyAlaHisIleAsnProAlaValThrValAlaCysLeuValGlyCysHisValSerPhe 83
```







```
Query Match: 25.72% Indels: 18
DB: 2 Gaps: 4
US-10-723-180-1 (1-1152) x US-08-393-996A-17 (1-269)
QY 139 TTCTGGAAGCAGTCACAGCGAATTTCTGGCCATCTTATTTTGTCTCTCCTCAGCCCTG 198
Db 10 PheTTPArgAlaValValAlaGluPheLeuAlaThrThrLeuPheValPheIleSerIle 29
QY 199 CGATCCACCATCAACTGGGGT-----GGAACAGAAAAGCCTTTACCGTCCGACATG 249
Db 30 GlySerAlaLeuGlyPheIleTyrProValGlyAsnAsnGlnThrAlaValGlnAspAsn 49
QY 250 GTTCTCATCTCCCTTTTGGACTCAGCATTTGCAACCATCGTGCGTCTTGGCCAT 309
Db 50 ValIysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis 69
QY 310 ATCAGCGGTGGCCACATCAACCTCGCAGTACTGTGGCCATCGTGTCACACGAGGATC 369
Db 70 IleSerGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeuSerCysGlnIle 89
QY 370 AGCATCGCCAGTCTCTCTTACATCGCAGCCGCTGCTGGGGCCATCATTCGAGCA 429
Db 90 SerIlePheArgAlaLeuMetTyrIleIleAlaGlnCysValGlyAlaIleValAlaThr 109
QY 430 GGAATCCTCTATCTGTGTACACCTCCAGTGTGTGGAGGCTGGGAGTCCACCATGTT 489
Db 110 AlaIleLeuSerGlyIleThrSerSerLeuThrGlyAsnSerLeuGlyArgAsnAspLeu 129
QY 490 CATGGAATCTTACCGCTGGTCATGCTCTCTGGTGTGATTAATCAATTCATTCATATG 549
Db 130 AlaAspGlyValAsnSerGlyGlnGlyLeuGlyIleGluIleIleGlyThrLeuGlnLeu 149
QY 550 GNGTTTACTCTTTGCCAGCTGTGATTCACACCGACTGTACTGCTCTGCTCAATAGCT 609
Db 150 ValLeuCysValLeuAlaThrThrAspArgArgArgAspLeuGlySerAlaPro 169
QY 610 TTAGCAATTTGGATTTCTGTGCAATTTGACATTTATTTTGCATCAATATATCTGTGCC 669
Db 170 LeuAlaIleGlyLeuSerValAlaLeuGlyHisLeuLeuAlaIleAspTyrThrGlyCys 189
QY 670 AGCATGAATCCGCGCATCTTTGGACCTCGCAGTATCATCGGAAATTTGGGAAACCAT 729
Db 190 GlyIleAsnProAlaArgSerPheGlySerAlaValIleThrHisAsnPheSerAsnHis 209
QY 730 TGGATATATTGGTTGGGCCCCATCATAGGAGTGTCTCGTGGTGGGCTTTATGAGTAT 789
Db 210 TrpIlePheTrpValGlyProPheIleGlyAlaLeuAlaValLeuIleIleTyrAspPhe 229
QY 790 GTCTTCTGTCAGATGTTGAATTCAAACGTGTTTAAAGAGCCCTTCAGCAAACTGCC 849
Db 230 IleLeuAlaPro-----ArgSerSerAspLeuThrAspArgValIys 243
QY 850 CAGCAAAACAAAAGGAGCTAC-----ATGGAGGTGGAGGACAAACAGGAGTACAG 897
Db 244 ValTTPThrSerGlyGlnValGluGluTyrAspLeuAspAlaAspIleAsnSerArg 263
QY 898 GTAGAGACGGATGACCTGATTTCTAAACCT 927
Db 264 ValGlu-----MetIysPro 268

RESULT 12
US-09-949-016-5934
; Sequence 5934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5934
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-5934

Alignment Scores:
Pred. No.: 1,548-51 Length: 269
Score: 561.00 Matches: 118
Percent Similarity: 60.74% Conservative: 46
Best local Similarity: 43.70% Mismatches: 88
Query Match: 25.72% Indels: 18
DB: 4 Gaps: 4
```

```
US-10-723-180-1 (1-1152) x US-09-949-016-5934 (1-269)
QY 139 TTCTGGAAGCAGTCACAGCGAATTTCTGGCCATCTTATTTTGTCTCTCCTCAGCCCTG 198
Db 10 PheTTPArgAlaValValAlaGluPheLeuAlaThrThrLeuPheValPheIleSerIle 29
QY 199 CGATCCACCATCAACTGGGGT-----GGAACAGAAAAGCCTTTACCGTCCGACATG 249
Db 30 GlySerAlaLeuGlyPheIleTyrProValGlyAsnAsnGlnThrAlaValGlnAspAsn 49
QY 250 GTTCTCATCTCCCTTTTGGACTCAGCATTTGCAACCATCGTGCGTCTTGGCCAT 309
Db 50 ValIysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis 69
QY 310 ATCAGCGGTGGCCACATCAACCTCGCAGTACTGTGGCCATCGTGTCACACGAGGATC 369
Db 70 IleSerGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeuSerCysGlnIle 89
QY 370 AGCATCGCCAGTCTGTCTTACATCGCAGCCGCTGCTGGGGCCATCATTCGAGCA 429
Db 90 SerIlePheArgAlaLeuMetTyrIleIleAlaGlnCysValGlyAlaIleValAlaThr 109
QY 430 GGAATCCTCTATCTGTGTACACCTCCAGTGTGTGGAGGCTGGGAGTCCACCATGTT 489
Db 110 AlaIleLeuSerGlyIleThrSerSerLeuThrGlyAsnSerLeuGlyArgAsnAspLeu 129
QY 490 CATGGAATCTTACCGCTGGTCTCTCTGGTGTGATTAATCAATTCATTCATATG 549
Db 130 AlaAspGlyValAsnSerGlyGlnGlyLeuGlyIleGluIleIleGlyThrLeuGlnLeu 149
QY 550 GTGTTTACTATCTTTGCCAGCTGTGATTCACACCGACTGTACTGCTCTGCTCAATAGCT 609
Db 150 ValLeuCysValLeuAlaThrThrAspArgArgArgAspLeuGlySerAlaPro 169
QY 610 TTAGCAATTTGGATTTCTGTGCAATTTGACATTTATTTTGCATCAATATATCTGTGCC 669
Db 170 LeuAlaIleGlyLeuSerValAlaLeuGlyHisLeuLeuAlaIleAspTyrThrGlyCys 189
QY 670 AGCATGAATCCGCGCATCTTTGGACCTCGCAGTATCATCGGAAATTTGGGAAACCAT 729
Db 190 GlyIleAsnProAlaArgSerPheGlySerAlaValIleThrHisAsnPheSerAsnHis 209
QY 730 TGGATATATTGGTTGGGCCCCATCATAGGAGTGTCTCGTGGTGGGCTTTATGAGTAT 789
Db 210 TrpIlePheTrpValGlyProPheIleGlyAlaLeuAlaValLeuIleIleTyrAspPhe 229
QY 790 GTCTTCTGTCAGATGTTGAATTCAAACGTGTTTAAAGAGCCCTTCAGCAAACTGCC 849
Db 230 IleLeuAlaPro-----ArgSerSerAspLeuThrAspArgValIys 243
QY 850 CAGCAAAACAAAAGGAGCTAC-----ATGGAGGTGGAGGACAAACAGGAGTACAG 897
```

Db 244 ValTprThrSerGlyGlnValGluGluTyrAspLeuAlaAspAlaAspIleAsnSerArg 263  
QY 898 GTAGACGGATGACCTGATTTCTAAACCT 927  
Db 264 ValGlu-----MetLysPro 268  
RESULT 13  
US-09-949-016-8317  
; Sequence 8317, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8317  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-8317

- Alignment Scores:  
Pred. No.: 281  
Score: 561.00 Matches: 118  
Percent Similarity: 60.74% Conservative: 46  
Best Local Similarity: 43.70% Mismatches: 88  
Query Match: 25.72% Indels: 18  
DB: 4 Gaps: 4

US-10-723-180-1 (1-1152) x US-09-949-016-8317 (1-281)

QY 139 TTCTGGAAGCAGTCACACGGAAATTCCTGGCCATGCTTATTTTGTCTCTCCTCAGCCTG 198  
Db 22 PheTprArgAlaValValAlaGluPheLeuAlaThrLeuPheValPheIleSerIle 41  
QY 199 GGATCCACCATCAACTGGGGT-----GGACAGAAAGCCTTTACCGTCGCACATG 249  
Db 42 GlySerAlaLeuGlyPheLysTyrProValGlyAsnAsnGlnThrAlaValGlnAspAsn 61  
QY 250 GTTCTCATCTCCCTTGTCTTGGACTCAGCATTCGACCATGGTCAGTCTTGGCCAT 309  
Db 62 ValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis 81  
QY 310 ATCAGCGGTGGCCACATCAACCTGCAGTACTGTGGCCATGTGTGCACACAGGAAGATC 369  
Db 82 IleSerGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeuSerCysGlnIle 101  
QY 370 AGCATGCCAAGTCTGTCTTACATCGCAGCCCATGCTGGGGGCCCATCATTCAGACA 429  
Db 102 SerIlePheArgAlaLeuMetTyrIleAlaGlnCysValGlyAlaIleValAlaThr 121  
QY 430 GGAATCCTATCTGGTCACACTCCCATGTGGTGGGAGGCGCTGGAGTCACCATGTT 489  
Db 122 AlaIleLeuSerGlyIleThrSerSerLeuThrGlyAsnSerLeuGlyArgAsnAspLeu 141  
QY 490 CATGAAATCTTACCGCTCATGCTCTCTGCTCTGCTGTAGTTGATATCATATTCATTTG 549  
Db 142 AlaAspGlyValAsnSerGlyGlnGlyLeuGlyIleGluIleIleGlyThrLeuGlnLeu 161  
QY 550 GTGTTTACTATCTTGGCCAGCTGTGATTCCAAACGACTGTGTGCATCGCTCAATAGCT 609  
Db 162 ValLeuCysValLeuAlaThrThrAspArgArgArgAspLeuGlyGlySerAlaPro 181

QY 610 TTACCAATTGATTTCTGTGTGCAATTGGACATTTATTTGCAATCAATTATCTACTGTCGC 669  
Db 182 LeuAlaIleGlyLeuSerValAlaLeuGlyHisLeuAlaIleAspTyrThrGlyCys 201  
QY 670 AGCATGAATCCCGCCGATCCTTTGGACCTGATCATGGAAATTTGGGAAACCAT 729  
Db 202 GlyIleAsnProAlaArgSerPheGlySerAlaValIleThrHisAsnPheSerAsnHis 221  
QY 730 TGGATATATTGGGTTGGCCCATCATAGGAGCTGTCTCGCTGGCTTTATGAGTAT 789  
Db 222 TrpIlePheTprValGlyProPheIleGlyAlaLeuAlaValLeuIleTyrAspPhe 241  
QY 790 GTCTTCTGTCAGATGTTGAATTCAAACGTCGTTTAAAGAGCCTTCAGCAAGCTGCC 849  
Db 242 IleLeuAlaPro-----ArgSerSerAspLeuThrAspArgVallys 255  
QY 850 CAGCAACAAAGGAAGCTAC-----ATGGAGGTGGAGGACCAACAGGAGTCAG 897  
Db 256 ValTprThrSerGlyGlnValGluGluTyrAspLeuAlaAspAlaAspIleAsnSerArg 275  
QY 898 GTAGACGGATGACCTGATTTCTAAACCT 927  
Db 276 ValGlu-----MetLysPro 280  
RESULT 14  
US-09-949-016-10471  
; Sequence 10471, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10471  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-10471  
Alignment Scores:  
Pred. No.: 1.57e-51 Length: 281  
Score: 561.00 Matches: 118  
Percent Similarity: 60.74% Conservative: 46  
Best Local Similarity: 43.70% Mismatches: 88  
Query Match: 25.72% Indels: 18  
DB: 4 Gaps: 4  
US-10-723-180-1 (1-1152) x US-09-949-016-10471 (1-281)  
QY 139 TTCTGGAAGCAGTCACACGGAAATTCCTGGCCATGCTTATTTTGTCTCTCAGCCTG 198  
Db 22 PheTprArgAlaValValAlaGluPheLeuAlaThrLeuPheValPheIleSerIle 41  
QY 199 GGATCCACCATCAACTGGGGT-----GGACAGAAAGCCTTTACCGTCGCACATG 249  
Db 42 GlySerAlaLeuGlyPheLysTyrProValGlyAsnAsnGlnThrAlaValGlnAspAsn 61  
QY 250 GTTCTCATCTCCCTTGTCTTGGACTCAGCATTCGACCATGGTCAGTCTTGGCCAT 309  
Db 62 ValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis 81  
QY 310 ATCAGCGGTGGCCACATCAACCTGCAGTACTGTGGCCATGTGTGCACACAGGAAGATC 369  
Db 82 IleSerGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeuSerCysGlnIle 101  
QY 370 AGCATGCCAAGTCTGTCTTACATCGCAGCCCATGCTGGGGGCCCATCATTCAGACA 429  
Db 102 SerIlePheArgAlaLeuMetTyrIleAlaGlnCysValGlyAlaIleValAlaThr 121  
QY 430 GGAATCCTATCTGGTCACACTCCCATGTGGTGGGAGGCGCTGGAGTCACCATGTT 489  
Db 122 AlaIleLeuSerGlyIleThrSerSerLeuThrGlyAsnSerLeuGlyArgAsnAspLeu 141  
QY 490 CATGAAATCTTACCGCTCATGCTCTCTGCTCTGCTGTAGTTGATATCATATTCATTTG 549  
Db 142 AlaAspGlyValAsnSerGlyGlnGlyLeuGlyIleGluIleIleGlyThrLeuGlnLeu 161  
QY 550 GTGTTTACTATCTTGGCCAGCTGTGATTCCAAACGACTGTGTGCATCGCTCAATAGCT 609  
Db 162 ValLeuCysValLeuAlaThrThrAspArgArgArgAspLeuGlyGlySerAlaPro 181

```
Db 82 IleSerGlyAlaHisLeuAenProAlaValThrLeuGlyLeuLeuSerCysGlnIle 101
QY 370 AGCATGCCAAGTCTGCTCTTACATCGCAGCCAGTCTGGGGCCCATCATTTGGAGCA 429
Db 102 SerIlePheArgAlaLeuMetTyrIleAlaGlnCysValGlyAlaIleValAlaThr 121
QY 430 GGAATCTCTATCTGTGTCACTCCAGTGTGGTGGAGGCTGGAGTCCACCATGTT 489
Db 122 AlaIleLeuSerGlyIleThrSerLeuThrGlyAenSerLeuGlyArgAenAspLeu 141
QY 490 CATGGAATCTTACCGTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 549
Db 142 AlaAspGlyValAenSerGlyGlnGlyLeuGlyIleGluIleIleGlyThrLeuGlnLeu 161
QY 550 GTGTTTACTATCTTCCAGCTGTGATTCACAAACGACCTGATCTCACTGGCTCAATAGCT 609
Db 162 ValLeuCysValLeuAlaThrThrAspArgArgAspLeuGlyGlySerAlaPro 181
QY 610 TTAGCAATTTGGATTTCTGTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTG 669
Db 182 LeuAlaIleGlyLeuSerValAlaLeuGlyHisLeuLeuAlaIleAspTyrThrGlyCys 201
QY 670 AGCATGAATCCCGCCGCTCTTTCGACCTGACCTGACCTGACCTGACCTGACCTGAC 729
Db 202 GlyIleAenProAlaArgSerPheGlySerAlaValIleThrHisAenPheSerAenHis 221
QY 730 TGGATATATTGGTGGGCGCCATCATAGGAGCTGCTCTGCTGCTGCTGCTGCTGCTG 789
Db 222 TrpIlePheTrpValGlyProPheIleGlyAlaLeuAlaValLeuIleTyrAspPhe 241
QY 790 GTCTCTCTGTCAGATGTTGAATTCAAACGCTGTTTAAAGAACCTTCAGCAAGCTGCC 849
Db 242 IleLeuAlaPro-----ArgSerSerAspLeuThrAspArgValLys 255
QY 850 CAGCAAAACAAAGGAAGCTAC-----ATGGAGGTGGAGGACAAACAGGAGTCAG 897
Db 256 ValTrpThrSerGlyGlnValGluGluThrAspLeuAspAlaAspIleAenSerArg 275
QY 898 GTAGAGACGATGACCTGATTCATAAACCT 927
Db 276 ValGlu-----MetLysPro 280

RESULT 15
US-08-447-554-5
; Sequence 5, Application US/08447554
; Patent No. 5661003
; GENERAL INFORMATION:
; APPLICANT: FUSHIMI, KIYOHIDE
; APPLICANT: UCHIDA, SHINICHI
; APPLICANT: SASAKI, SEI
; APPLICANT: MARUMO, FUMIAKI
; TITLE OF INVENTION: WATER CHANNEL
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. NW, Ste. 5500
; CITY: Washington, DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,554
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,365
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
```

```
;
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 5100-0003.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 269 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-447-554-5

Alignment Scores:
Pred. No.: 1,97e-51 Length: 269
Score: 560.00 Matches: 118
Percent Similarity: 60.74% Conservative: 46
Best Local Similarity: 43.70% Mismatches: 88
Query Match: 25.68% Indels: 18
DB: 1 Gaps: 4

US-10-723-180-1 (1-1152) x US-08-447-554-5 (1-269)
QY 139 TTCTGGAAGCAGTCACAGCGAATTTCTGGCCATGCTTATTTTGTCTCTCTCAGCCTG 198
Db 10 PheTrpArgAlaValAlaGluPheLeuAlaThrThrLeuPheValPheIleSerIle 29
QY 199 GGATCCACCATCACTGGGTT-----GGACAGAAAGCCCTTTACCGTCCACATG 249
Db 30 GlySerAlaLeuGlyPheLysTyrProValGlyAenAenGlnThrAlaValGlnAspAen 49
QY 250 GTTCTCATCTCCCTTTGCTTTGGACTCAGCATTTGCAACCATGTCAGTGTCTTGGCCAT 309
Db 50 ValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis 69
QY 310 ATCAGCGTGGCCACATCAACCTGCACTGCTGTCAGTGTGGCCATGTCACCAAGGAGATC 369
Db 70 IleSerGlyAlaHisLeuAenProAlaValThrLeuGlyLeuLeuSerCysGlnIle 89
QY 370 AGCATGCCAAGTCTGCTCTTACATCGACCCAGTCTCTGGGGCCCATCATTTGGAGCA 429
Db 90 SerIlePheArgAlaLeuMetTyrIleIleAlaGlnCysValGlyAlaIleValAlaThr 109
QY 430 GGAATCTCTATCTGTGTCCACACCTCCAGTGTGGGAGGCGCTGGAGTCCACCATGTT 489
Db 110 AlaIleLeuSerGlyIleThrSerSerLeuThrGlyAenSerLeuGlyArgAenAspLeu 129
QY 490 CATGGAATCTTACCGTGTGTCATGCTCTCTGTTGAGTTGATTAATCACAATTTCAATTG 549
Db 130 AlaAspGlyValAenSerGlyGlnGlyLeuGlyIleGluIleIleGlyThrLeuGlnLeu 149
QY 550 GTGTTTACTATCTTTCAGCTGTGATTCACAAAGGACTGATGTCACCTGGCTCAATAGCT 609
Db 150 ValLeuCysValLeuAlaThrThrAspArgArgArgAspLeuGlyGlySerAlaPro 169
QY 610 TTAGCAATTTGGATTTCTGTTGCAATTTATTTGCAATTTATTTCAATTTACTGTCGCC 669
Db 170 LeuAlaIleGlyLeuSerValAlaLeuGlyHisLeuLeuAlaIleAspTyrThrGlyCys 189
QY 670 AGCATGAATCCCGCCGCTCTTTCGACCTGACCTGACCTGACCTGACCTGACCTGAC 729
Db 190 GlyIleAenProAlaArgSerPheGlySerAlaValIleThrHisAenPheSerAenHis 209
QY 730 TGGATATATTGGTGGGCGCCATCATAGGAGCTGCTCTGCTGCTGCTGCTGCTGCTGCT 789
Db 210 TrpIlePheTrpValGlyProPheIleGlyAlaLeuAlaValLeuIleTyrAspPhe 229
QY 790 GTCTCTCTGTCAGATGTTGAATTCAAACGCTGTTTAAAGAACCTTCAGCAAGCTGCC 849
Db 230 IleLeuAlaPro-----ArgSerSerAspLeuThrAspArgValAen 243
QY 850 CAGCAAAACAAAGGAAGCTAC-----ATGGAGGTGGAGGACAAACAGGAGTCAG 897
```

```

Db      244 ValTrpThrSerGlyGlnValGluGluTyrAspLeuAspAlaAspAspIleAsnSerArg 263
          ||| |||          :||| |||
QY      898 GTAGAGACGATGACCTGATTCTAAACCT 927
          ||| |||          :||| |||
Db      264 ValGlu-----MetLysPro 268
          ||| |||          :||| |||

```

Search completed: September 8, 2005, 12:16:57  
Job time : 48 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: September 8, 2005, 11:59:33 ; Search time 192 Seconds  
(without alignments)  
4725.879 Million cell updates/sec

Title: US-10-723-180-1  
Perfect score: 2181  
Sequence: 1 ggggcaggcaatgagagctg.....attactcagctcaacaata 1152

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 3548624

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p.model -DRV=xlp  
-O=/cgn2\_1/USPTO\_spool\_p/US10723180/runat\_08092005\_123257\_29276/app\_query.fasta\_1.1351  
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USR=US10723180 @CGN 1 1 199 @runat\_08092005\_123257\_29276  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=0.5  
-FGAPOF=6 -FGAPEXT=7 -YGAPOF=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pcp.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pcp.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pcp.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pcp.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pcp.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pcp.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pcp.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pcp.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pcp.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pcp.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pcp.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pcp.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pcp.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pcp.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pcp.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pcp.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pcp.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pcp.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pcp.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pcp.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pcp.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1685	77.3	323	15	US-10-295-027-254 Sequence 254, App
2	1634	74.9	323	17	US-10-732-923-15771 Sequence 15771, A
3	1628	74.6	323	17	US-10-732-923-15770 Sequence 15770, A
4	1623	74.4	323	17	US-10-732-923-15775 Sequence 15775, A
5	1574	72.2	323	17	US-10-732-923-15576 Sequence 15576, A
6	1567	71.8	323	17	US-10-732-923-15577 Sequence 15577, A
7	1542.5	70.7	322	17	US-10-732-923-15579 Sequence 15579, A
8	1524	69.9	355	17	US-10-732-923-15574 Sequence 15574, A
9	1516	69.5	301	17	US-10-732-923-15572 Sequence 15572, A
10	1499.5	68.8	354	17	US-10-732-923-15578 Sequence 15578, A
11	1483	68.0	300	17	US-10-732-923-15575 Sequence 15575, A
12	1458.5	66.9	300	17	US-10-732-923-15580 Sequence 15580, A
13	656.5	30.1	249	17	US-10-732-923-15746 Sequence 15746, A
14	625	28.7	251	17	US-10-732-923-15615 Sequence 15615, A
15	617	28.3	265	14	US-10-097-340-12 Sequence 12, Appl
16	617	28.3	265	14	US-10-171-311-16 Sequence 16, Appl
17	617	28.3	265	17	US-10-645-756-6 Sequence 6, Appli
18	617	28.3	265	17	US-10-756-149-4696 Sequence 4696, Ap
19	606	27.8	263	17	US-10-732-923-15773 Sequence 15773, A
20	604	27.7	273	17	US-10-732-923-15761 Sequence 15761, A
21	604	27.7	271	17	US-10-732-923-15583 Sequence 15583, A
22	602	27.6	262	17	US-10-732-923-15760 Sequence 15760, A
23	602	27.6	263	17	US-10-732-923-15763 Sequence 15763, A
24	601	27.6	271	17	US-10-732-923-15584 Sequence 15584, A
25	599	27.5	262	17	US-10-732-923-15765 Sequence 15765, A
26	597.5	27.4	263	17	US-10-732-923-15581 Sequence 15581, A
27	595.5	27.3	263	17	US-10-732-923-15582 Sequence 15582, A
28	595	27.3	274	17	US-10-732-923-15762 Sequence 15762, A
29	589	27.0	271	18	US-10-794-899-95 Sequence 95, Appl
30	561	25.7	269	16	US-10-476-728A-6 Sequence 6, Appli
31	530.5	24.3	251	17	US-10-732-923-15661 Sequence 15661, A
32	525.5	24.1	222	17	US-10-645-756-10 Sequence 10, Appl
33	515	23.6	255	17	US-10-732-923-15667 Sequence 15667, A
34	508	23.3	282	15	US-10-403-161-10 Sequence 10, Appl
35	496.5	22.8	239	17	US-10-732-923-15747 Sequence 15747, A
36	496.5	22.8	245	17	US-10-732-923-15748 Sequence 15748, A
37	449.5	20.6	271	17	US-10-732-923-15745 Sequence 15745, A
38	447	20.5	72	9	US-09-864-761-35778 Sequence 35778, A
39	440	20.2	167	15	US-10-188-186-50 Sequence 50, Appl
40	438	20.1	248	16	US-10-437-963-178122 Sequence 178122, A
41	437	20.0	290	16	US-10-437-963-156669 Sequence 156669, A
42	436.5	20.0	210	17	US-10-645-756-8 Sequence 8, Appli
43	435	19.9	248	16	US-10-767-701-46122 Sequence 46122, A
44	434	19.9	247	15	US-10-424-599-191248 Sequence 191248, A
45	434	19.9	247	17	US-10-732-923-15685 Sequence 15685, A

ALIGNMENTS

RESULT 1  
US-10-295-027-254  
; Sequence 254, Application US/10295027  
; Publication No. US20030232350A1  
; GENERAL INFORMATION:  
; APPLICANT: Afar, Daniel  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Ginsberg, Wendy M.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Glynn, Richard  
; APPLICANT: Hevezi, Peter A.  
; APPLICANT: Mack, David H.  
; APPLICANT: Murray, Richard  
; APPLICANT: Watson, Susan R.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer  
; FILE REFERENCE: 018501-01250005  
; CURRENT APPLICATION NUMBER: US/10/295,027  
; CURRENT FILING DATE: 2002-11-13



```
Db      41  GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGly 60
QY      220  GGAACAGAAAGCCTTTACGGTCGACATGGTTCTCATCTCCCTTTGCTTGGACTCAGC 279
Db      61  GlyAlaGluIlyBProleuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer 80
QY      280  ATTGCAACCATGGTCAGTCGCTTTGGCCATATCAGCGGTGGCCACATCAACCTCGAGTG 339
Db      81  IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100
QY      340  ACTGTGGCCATGGTGGACCGAGGAGATCAGCATCGCCAGTCGCTCTCTACATCGCA 399
Db      101  ThrValAlaMetValCysThrArgArgIleSerIleAlaLysSerValPheTyrIleAla 120
QY      400  GCCAGTGCCTGGGGCCCATCATTTGAGCAGGAAATCCTCTATCTGTGTCAACCTCCAGT 459
Db      121  AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValThrProPheSer 140
QY      460  GTGGTGGGAGGCTGGGAGTCACCATGTCATGGTTCATGGAAATCTTACCGCTCGTGTCTC 519
Db      141  ValValGlyGlyLeuGlyValThrValHisGlyAsnLeuSerAlaGlyHisGlyLeu 160
QY      520  CTGGTTGAGTTCATATCATATTCATTTCAATTTGGTTTACTATCTTTGCCAGCTGTGATTC 579
Db      161  LeuValGluLeuIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180
QY      580  AAACGGACTGATGTCTACTGGCTCAATAGCTTTAGCAATTTGGAAATTTCTGTGTGCAATTGGA 639
Db      181  LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyIleSerValAlaIleGly 200
QY      640  CATTTATTTGCAATCAATATATCTGTGTCAGCATGAATCCCGCCGATCCTTTGGACCT 699
Db      201  HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220
QY      700  GCAGTTATCATGGGAAATTTGGGAAACCATTTGGATATATTGGGTTGGGCCCATCATAGA 759
Db      221  AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly 240
QY      760  GCTGTCTCGCTGGTGGCTTTATAGTATGTCTCTGTCCAGATGTTGAATTCAAACGT 819
Db      241  AlaValLeuAlaGlyLeuTyrGluTyrValPheCysProAspValGluLeuLysArg 260
QY      820  CGTTTTAAAGAGCCTTCAGCAAGCTGCCAGCAACCAAAAGAGAGCTACATGGAGGTG 879
Db      261  ArgPheLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluVal 280
QY      880  GAGGACAAAGAGTTCAGTACGAGCGATGACCTGATTTCTTAAACCTGGAGTGGTGAT 939
Db      281  GluAspAsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGlyValValHis 300
QY      940  GTGATTGACGTTGACCGGGAGAGGAGAGAGAGGAGAGCAATCTCGAGAGGTATTG 999
Db      301  ValIleAspIleAspArgGlyGluGluLysLysGlyLysAspProSerGlyGluValLeu 320
QY      1000  TCTTCAGTA 1008
Db      321  SerSerVal 323
```

## RESULT 3

```
US-10-732-923-15770
; Sequence 15770, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15770
; LENGTH: 323
```

```
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-732-923-15770
Alignment Scores:
Pred. No.:      6,17e-149      Length:      323
Score:          1628.00      Matches:     311
Percent Similarity: 97.52%      Conservative: 4
Best Local Similarity: 96.28%      Mismatches: 8
Query Match:     74.64%      Indels:      0
DB:              17          Gaps:         0
```

US-10-723-180-1 (1-1152) x US-10-732-923-15770 (1-323)

```
QY      40  ATGAGTGCAGACCCACAGCAAGCGGTGGGTAAAGTGTGGACCTTTGTGTACCAGAGAG 99
Db      1  MetSerAspArgProAlaAlaThrArgTrpGlyLysCysGlyProleuCysThrArgGlu 20
QY      100  AACATCATGTGTGGCTTTCAAAGGGGTCTGGACTCAAGCTTTCTGGAAAGCAGTCACAGCG 159
Db      21  SerIleMetValAlaPheLysGlyValTrpThrGlnThrPheTrpLysAlaValThrAla 40
QY      160  GAATTTCTGGCCCATGCTTATTTTGTTCCTCCTCAGCTGGGATCCACCATCAACTGGGT 219
Db      41  GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGly 60
QY      220  GGAACAGAAAGCCTTTACCGGTCCACATGGTTCTCATCTCCCTTTGCTTGGACTCAGC 279
Db      61  GlyAlaGluLysProleuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer 80
QY      280  ATTGCAACCATGTGTCAGTGCCTTTTGGCCATATCAGCGGTGGCCACATCAACCTCGAGTG 339
Db      81  IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100
QY      340  ACTGTGGCCATGGTGTGCACGAGGAAGATCAGCATCGCCAAAGTCTGTCTTCTACATCGCA 399
Db      101  ThrValAlaMetValCysThrArgArgIleSerIleAlaLysSerValPheTyrIleAla 120
QY      400  GCCAGTGCCTGGGGCCCATCATTTGAGCAGGAAATCCTCTATCTGTGTCCAGCTCAGCTCC 459
Db      121  AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValThrProPheSer 140
QY      460  GTGGTGGGAGGCTGGGAGTCACCATGTCATGGTTCATGGAAATCTTACCGCTCGTGTCTC 519
Db      141  ValValGlyGlyLeuGlyValThrValHisGlyAsnLeuSerAlaGlyHisGlyLeu 160
QY      520  CTGGTTGAGTTCATATCATATTCATTTCAATTTGGTTTACTATCTTTGCCAGCTGTGATTC 579
Db      161  LeuValGluLeuIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180
QY      580  AAACGGACTGATGTCTACTGGCTCAATAGCTTTAGCAATTTGGAAATTTCTGTGTGCAATTGGA 639
Db      181  LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyIleSerValAlaIleGly 200
QY      640  CATTTATTTGCAATCAATATATCTGTGTCAGCATGAATCCCGCCGATCCTTTGGACCT 699
Db      201  HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220
QY      700  GCAGTTATCATGGGAAATTTGGGAAACCATTTGGATATATTGGGTTGGGCCCATCATAGA 759
Db      221  AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly 240
QY      760  GCTGTCTCGCTGGTGGCTTTATAGTATGTCTCTGTCCAGATGTTGAATTCAAACGT 819
Db      241  AlaValLeuAlaGlyLeuTyrGluTyrValPheCysProAspValGluLeuLysArg 260
QY      820  CGTTTTAAAGAGCCTTCAGCAAGCTGCCAGCAACCAAAAGAGAGCTACATGGAGGTG 879
Db      261  ArgPheLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluVal 280
QY      880  GAGGACAAAGAGTTCAGTACGAGCGATGACCTGATTTCTTAAACCTGGAGTGGTGAT 939
Db      281  GluAspAsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGlyValValHis 300
QY      940  GTGATTGACGTTGACCGGGAGAGGAGAGAGGAGAGCAATCTCGAGAGGTATTG 999
Db      301  ValIleAspIleAspArgGlyGluGluLysLysGlyLysAspProSerGlyGluValLeu 320
```



QY 220 GGAAACAGAAAGCTTTACCGGTGCGACATGGTTCTCATCTCCCTTTGCTTTGGACCTCAGC 279  
|||:|||||  
Db 61 GlySerGluAsnProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer 80  
|||:|||||  
QY 280 ATTGCAACCATGGTGGAGTCTTTGGCCATCATGAGGTGGCCACATCAACCTCGAGTG 339  
|||:|||||  
Db 81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100  
|||:|||||  
QY 340 ACTGTGGCCATGGTGTGCGACAGGAGATCAGCATCGCCAAAGTCTGTCTTCTACATCGCA 399  
|||:|||||  
Db 101 ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrlleIle 120  
|||:|||||  
QY 400 GCCCAGTGCCTGGGGGCCATCATTTGGAGCAGGAATCCTCTATCTGTGTCACACCTCCAGT 459  
|||:|||||  
Db 121 AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrlleValThrProSer 140  
|||:|||||  
QY 460 GTGGTGGGAGCCCTGGAGTGCACCATGGTTCATGGAATCTTACCGCTGTGATCTC 519  
|||:|||||  
Db 141 ValValIGlyLeuGlyValThrValHisGlyAsnLeuThrAlaGlyHisGlyLeu 160  
|||:|||||  
QY 520 CTGGTGTGAGTTGATTAATCAATTTCAATTTGCTGTGTACTATCTTTCGAGCTGTGATTC 579  
|||:|||||  
Db 161 LeuValGluLeuIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180  
|||:|||||  
QY 580 AACCGACTGATCTCACTGGCTCAATGAGCTTTAGCAATTTGGAATTTCTGTGCAATTGGA 639  
|||:|||||  
Db 181 LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGly 200  
|||:|||||  
QY 640 CATTATTGTCATCAATTAATCTATCTGTCAGCATGAATCCCGCCGATCCTTTGGACCT 699  
|||:|||||  
Db 201 HisLeuPheAlaIleAsnTyrlleValAlaSerMetAsnProAlaArgSerPheGlyPro 220  
|||:|||||  
QY 700 GCAGTTTATCATCGGAAATTTGGAAACCATTTGGATATATTGGTGGCCCATCATAGGA 759  
|||:|||||  
Db 221 AlaValIleMetGlyAsnTrpAlaAsnHisTrpIleTyrlleValGlyProIleMetGly 240  
|||:|||||  
QY 760 GCTGTCTCGTGGTGGCTTTATGAGTAGTGTCTTCTGTCAGATGTGAATTTCAACCT 819  
|||:|||||  
Db 241 AlaValIleMetGlyAsnTrpAlaAsnHisTrpIleTyrlleValGlyProIleMetGly 260  
|||:|||||  
QY 820 CGTTTTAAAGAGCCCTTACGAAAGCTGCCAGCAAAACAAAGAGAGCTACATGGAGGTG 879  
|||:|||||  
Db 261 ArgLeuLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrlleMetGluVal 280  
|||:|||||  
QY 880 GAGGACACAGGAGTCAAGTAGAGCGATGACCTGATTTCTTAAACCTGGAGTGTGCTAT 939  
|||:|||||  
Db 281 GluAspAsnArgSerGlnValGluThrGluAspLeuIleLeuLysProGlyValValHis 300  
|||:|||||

## RESULT 6

US-10-732-923-15577  
; Sequence 15577, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732, 923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 15577  
; LENGTH: 323  
; TYPE: PR

; ORGANISM: Mus musculus  
US-10-732-923-15577  
Alignment Scores:  
Pred. No.: 5,26e-143 Length: 323  
Score: 1567.00 Matches: 300  
Percent Similarity: 95.36% Conservative: 15  
Best Local Similarity: 92.88% Mismatches: 8  
Query Match: 71.85% Indels: 0  
DB: 17 Gaps: 0

US-10-723-180-1 (1-1152) x US-10-732-923-15577 (1-323)

QY 40 ATGAGTGCACAGACCCACAGCAAGGGCGGTGAGTGGACCTTTGTGTACACAGAG 99  
|||:|||||  
Db 1 MetSerAspGlyAlaAlaAlaArgTrpGlyLysCysGlyHisSerCysSerArgGlu 20  
|||:|||||  
QY 100 AACATCATGGTGGCTTTCAAAGGGGTCTGAGCTCAAGCTTTTGGAAAGAGTCAACGCG 159  
|||:|||||  
Db 21 SerIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValSerAla 40  
|||:|||||  
QY 160 GAATTTCTGGCCATGCTTATTTTCTCTCCTCAGCTGGGATCCACCATCACTGGGT 219  
|||:|||||  
Db 41 GluPheLeuAlaThrLeuIlePheValLeuLeuGlyValGlySerThrIleAsnTrpGly 60  
|||:|||||  
QY 220 GGAAACAGAAAGCTTTACCGGTGCGACATGGTTCTCATCTCCCTTTGCTTTGGACCTCAGC 279  
|||:|||||  
Db 61 GlySerGluAsnProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer 80  
|||:|||||  
QY 280 ATTGCAACCATGGTGGAGTCTTTGGCCATCATGAGGTGGCCACATCAACCTCGAGTG 339  
|||:|||||  
Db 81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100  
|||:|||||  
QY 340 ACTGTGGCCATGGTGTGCGACAGGAGATCAGCATCGCCAAAGTCTGTCTTACATCGCA 399  
|||:|||||  
Db 101 ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrlleIle 120  
|||:|||||  
QY 400 GCCCAGTGCCTGGGGCCATCATTTGGAGCAGGAATCCTCTATCTGTCACACCTCCACGT 459  
|||:|||||  
Db 121 AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrlleValThrProSer 140  
|||:|||||  
QY 460 GTGGTGGGAGGCTGGAGTCCACATGGTTCATGGAATTTTACCGCTGGTCAATGCTCTC 519  
|||:|||||  
Db 141 ValValGlyGlyLeuGlyValThrValHisGlyAsnLeuThrAlaGlyHisGlyLeu 160  
|||:|||||  
QY 520 CTGGTGTGAGTTGATTAATCAATTTCAATTTGCTGTGTACTATCTTTCGAGCTGTGATTC 579  
|||:|||||  
Db 161 LeuValGluLeuIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180  
|||:|||||  
QY 580 AACCGACTGATGTCACTGGCTCAATGAGCTTTAGCAATTTGGAATTTCTGTGCAATTGGA 639  
|||:|||||  
Db 181 LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGly 200  
|||:|||||  
QY 640 CATTATTGTCATCAATTAATCTATCTGTCAGCATGAATCCCGCCGATCCTTTGGACCT 699  
|||:|||||  
Db 201 HisLeuPheAlaIleAsnTyrlleValAlaSerMetAsnProAlaArgSerPheGlyPro 220  
|||:|||||  
QY 700 GCAGTTTATCATCGGAAATTTGGAAACCATTTGGATATATTGGTGGCCCATCATAGGA 759  
|||:|||||  
Db 221 AlaValIleMetGlyAsnTrpAlaAsnHisTrpIleTyrlleValGlyProIleMetGly 240  
|||:|||||  
QY 760 GCTGTCTCGTGGTGGCTTTATGAGTAGTGTCTTCTGTCAGATGTGAATTTCAACCT 819  
|||:|||||  
Db 241 AlaValLeuAlaGlyAlaLeuTyrlleValPheCysProAspValGluLeuLysArg 260  
|||:|||||  
QY 820 CGTTTTAAAGAGCCCTTACGAAAGCTGCCAGCAAAACAAAGAGAGCTACATGGAGGTG 879  
|||:|||||  
Db 261 ArgLeuLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrlleMetGluVal 280  
|||:|||||  
QY 880 GAGGACACAGGAGTCAAGTAGAGCGATGACCTGATTTCTTAAACCTGGAGTGTGCTAT 939  
|||:|||||  
Db 281 GluAspAsnArgSerGlnValGluThrGluAspLeuIleLeuLysProGlyValValHis 300  
|||:|||||





Qy	253	CTCATCTCCCTTTGCTTTGGACTCAGCATTCGCAACCATGGTCAGTGTCTTTGGCCATATTC	312
Db	104	LeuIleSerLeuCyAPheGlyLeuSerIleAlaThrMetValGlnCysPheGlyHisIle	123
Qy	313	AGCGGTGGCCACATCAACCCCTCAGTGTACTGTGGCCATGGTGTGCACCAGGAAGATCAGC	372
Db	124	SerGlyGlyHisIleAsnProAlaValThrValAlaMetValCysThrArgLysIleSer	143
Qy	373	ATCGCCAAAGTCGTCTTCTACATCCGACGCCCATGTGCCTGGGGGCCCATCATTTGGAGCAGA	432
Db	144	IleAlaLysSerValPheTyrIleIleAlaGlnCysAlaLeuGlyAlaIleIleGlyAlaGly	163
Qy	433	ATCCTCTATCTGGTCACACCTCCAGTGTGGTGGAGGCGCTGGGAGTACCACATGGTTTCAT	492
Db	164	IleLeuTyrLeuValThrProSerValValGlyLeuGlyValThrThrValHis	183
Qy	493	GGMAATCTTACCGCTGGTCATCGTCTCTCTGGTTGAGTTGATATACATATTCATTTCAATGGTG	552
Db	184	GlyAsnLeuThrAlaGlyHisGlyLeuLeuValGluLeuIleIleThrPheGlnLeuVal	203
Qy	553	TTTACTATCTTTGGCCAGCTGTGATTCCAAACGGAGCTGATGTCACTGGCTCAATAGCTTTTA	612
Db	204	PheThrIlePheAlaSerCysAspSerLysArgThrAspValThrGlySerIleAlaLeu	223
Qy	613	GCAATTTGGAATTTCTGTTCGAATTGGACATTTATTTTGGCAATCAATATATCTCGTGCCACG	672
Db	224	AlaIleGlyPheSerValAlaIleGlyHisLeuPheAlaIleAsnTyrThrGlyAlaSer	243
Qy	673	ATGAATCCCGCCGATCCTTTGGACCTCGAGTTATCATGGGAAATTTGGGAAACCATTTGG	732
Db	244	MetAsnProAlaArgSerPheGlyProAlaValIleMetGlyAsnTrpAlaAsnHisTrp	263
Qy	733	ATATATTGGGTGGGCCCATCATAGAGCTGTTCCTCGTGTGGTGGCTTTATCAGTATATGTC	792
Db	264	IleTyrTrpValGlyProIleMetGlyAlaValLeuAlaGlyAlaLeuTyrGluTyrVal	283
Qy	793	TTCTGTCAGATGTTGAAATTCAAACGTGCTTTTTAAAGAACCTTCAGCAAGCTCCCCAG	852
Db	284	PheCysProAspValGluLeuLysArgLeuLysGluAlaPheSerLysAlaAlaGln	303
Qy	853	CAAAACAAAGGAAGCTACATGGAGTGGGGAACACAGGAGTCCAGGTAGACGATGATC	912
Db	304	GlnThrLysGlySerTyrMetGluValGluAspAsnArgSerGlnValGluThrGluAsp	323
Qy	913	CTGATTTCTAAAAACCTGGAGTGGTGCATGTGATTGACGTGTACCGGGGAGAGAGGAAG	972
Db	324	LeuIleLeuLysProGlyValValHisValIleAspIleAspArgGlyGluGluLysLys	343
Qy	973	GGGAAAGACCAATCTGGAGAGGTATTGTCTTCAGTA	1008
Db	344	GlyLysAspSerSerGlyGluValLeuSerSerVal	355

RESULT 9		
US-10-732-923-15772		
Sequence 15772, Application US/10732923		
Publication No. US20050108791A1		
GENERAL INFORMATION:		
APPLICANT: Edgerton, Michael D		
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES		
FILE REFERENCE: 39-15(52796)C		
CURRENT APPLICATION NUMBER: US/10/732,923		
CURRENT FILING DATE: 2003-12-10		
PRIOR APPLICATION NUMBER: 10/310,154		
PRIOR FILING DATE: 2002-12-04		
NUMBER OF SEQ ID NOS: 24149		
SEQ ID NO 15772		
LENGTH: 301		
TYPE: PRT		
ORGANISM: Bos taurus		
US-10-732-923-15772		
Alignment Scores:	4.65e-138	Length:
Pred. No.:		



```
US-10-732-923-15578
; Sequence 15578, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15578
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-732-923-15578

Alignment Scores:
Pred. No.: 1,98e-136 Length: 354
Score: 1499.50 Matches: 288
Percent Similarity: 95.51% Conservatative: 10
Best Local Similarity: 92.31% Mismatches: 13
Query Match: 68.75% Indels: 1
DB: 17 Gaps: 1

US-10-723-180-1 (1-1152) x US-10-732-923-15578 (1-354)
QY 73 AAGTGTGACCTTGTGTACAGAGACATCATGGTGGCTTCAAAGGGGTCTGGACT 132
DB 44 LysCysGlyHisSerCysSerArgGluSerIleMetValAlaPheLysGlyValThr 63
QY 133 CAAGCTTCTGAAAGCAGTCAACGGAATTTCTGGCAGTCTATTTTGTCTCCTC 192
DB 64 GlnAlaPheTrpLysAlaValSerAlaGluPheLeuAlaThrLeuIlePheVal---Leu 82
QY 193 AGCCCTGGATCCACCATCAACCTGGGTGGAAACAGAAAGCCCTTACCGGTGCACATGTT 252
DB 83 GlyValGlySerThrIleAsnTrpGlyGlySerGluAsnProLeuProValAspMetVal 102
QY 253 CTCATCTCCCTTGTGTTGGACTCAGCATTCGAACCATGGTGCAGTGTCTGGCCATATC 312
DB 103 LeuIleSerLeuCysPheGlyLeuSerIleAlaThrMetValGlnCysLeuGlyHisIle 122
QY 313 ACCGTGGCCACATCAACCTCCAGTCACTGTGGCCATGGTGTGCACAGGAAGATCAGC 372
DB 123 SerGlyGlyHisIleAsnProAlaValThrValAlaMetValCysThrArgLysIleSer 142
QY 373 ATCGCCAAAGTCTGTCTTACATCCAGCCAGTCCCTGGGGCCATCATTTGGAGCAGA 432
DB 143 IleAlaLysSerValPheTrpIleIleAlaGlnCysLeuGlyAlaIleIleGlyAlaGly 162
QY 433 ATCTCTATCTCGTCCACACTCCCACTGTGGTGGAGGCTGGAGTCAACCATGGTTCAT 492
DB 163 IleLeuTrpLeuValThrProSerValValGlyGlyLeuGlyValThrThrValHis 182
QY 493 GGAATCTTACCGTGGTCACTGCTCTCGTGTGATGATTAATCACATTTCAATTTGGTG 552
DB 183 GlyAsnLeuThrAlaGlyHisGlyLeuValGluLeuIleIleThrPheGlnLeuVal 202
QY 553 TTTACTATCTTTCGAGCTGATTCCTCAACCGACTGATGTACCTGGCTCAATAGCTTTA 612
DB 203 PheThrValPheAlaSerCysAspSerLysArgThrAspValThrGlySerIleAlaLeu 222
QY 613 GCAATTGGAATTTCTGTGCAATTTGACATTTATTTGCAATCAATATATCTGGTCCAGC 672
DB 223 AlaIleGlyPheSerValAlaIleGlyHisLeuPheAlaIleAsnTrpThrGlyAlaSer 242
QY 673 ATGAATCCCGCCGATCTCTTGGACCTGAGTATATCATGGGAATTTGGAAACCATGG 732
DB 243 MetAsnProAlaArgSerPheGlyProAlaValIleMetGlyAsnTrpAlaAsnHisTrp 262
QY 733 ATATATTGGGTTGGGCCCATCATAGGAGCTGTCTCGCTGGTGGCTTTATGAGTATGTC 792
```

---

```
Db 263 IleTrpValGlyProIleMetGlyAlaValLeuAlaGlyAlaLeuTrpGluTrpVal 282
QY 793 TTCTGTCAGATGTTGAATTCAAACGTCGTTTAAAGAACCTTTCAGCAAGCTGCCAG 852
DB 283 PheCysProAspValGluLeuLysArgGluLysGluAlaPheSerLysAlaGln 302
QY 853 CAACAAAGAAAGCTACATCGAGGTGGAGACAAACAGGAGTCAGGTAGACGGATGAC 912
DB 303 GlnThrLysGlySerTrpMetGluValGluAspAsnArgSerGlnValGluThrGluAsp 322
QY 913 CTGATTCTAAACCTGGAGTGGTGCATGTCATTCACGTTGACCGGGGAGAGGAGAAG 972
DB 323 LeuIleLeuLysProGlyValValHisValIleAspIleAspArgGlyGluLysLys 342
QY 973 GGGAAAGACCAATCTGGAGAGGTATTGCTTTCAGTA 1008
DB 343 GlyArgAspSerSerGlyGluValLeuSerSerVal 354

RESULT 11
US-10-732-923-15575
; Sequence 15575, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15575
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-732-923-15575

Alignment Scores:
Pred. No.: 7,51e-135 Length: 301
Score: 1483.00 Matches: 285
Percent Similarity: 96.68% Conservatative: 6
Best Local Similarity: 94.68% Mismatches: 10
Query Match: 68.00% Indels: 0
DB: 17 Gaps: 0

US-10-723-180-1 (1-1152) x US-10-732-923-15575 (1-301)
QY 106 ATGCTGGCTTTCAAAGGGTCTGGACTCAAGCTTCTGGAAAGCAGTCACAGCGGAATTT 165
DB 1 MetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValSerAlaGluPhe 20
QY 166 CTGGCCATGCTTATTTTGTCTCTCAGCTGGATCCACCATCAACTGGGTGGAAACA 225
DB 21 LeuAlaThrLeuIlePheValLeuLeuGlyValGlySerThrIleAsnTrpGlySer 40
QY 226 GAAAGCCTTTACCGGTGCACATGTTCTCATCTCCCTTTGCTTTGGACTCAGCATTCGA 285
DB 41 GluAsnProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAla 60
QY 286 ACCATGTTGAGTGTGTTGGCCATATCAGCGGTGGCCACATCAACCTTCGAGTGCATGTG 345
DB 61 ThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaValThrVal 80
QY 346 GCCATGTTGTGCACAGGAACATCAGCATCCCAAGTCTGTCTTCTATCATCGAGCCAGC 405
DB 81 AlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTrpIleIleAlaGln 100
QY 406 TGCCTGGGGGCCATCATTTGGAGCAGGAATCTCTATCTGTGTCACACCTCCAGGTGTGTG 465
DB 101 CysLeuGlyAlaIleIleGlyAlaGlyIleLeuTrpLeuValThrProProSerValVal 120
QY 466 GGAGCCCTGGGAGTCACCATGTTTCATGGAAATCTTACCGCTGTGTCATGGTCTCTCTGTT 525
```

Db 121 GlyGlyLeuGlyValThrThrValHisGlyAsnLeuThrAlaGlyHisGlyLeuLeuVal 140  
QY 526 GAGTTGATATCACATTTCAATTTGGTGTCTTACTATCTTTCCAGCTGTGATTTCCAAACGG 585  
Db 141 GluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSerLysArg 160  
QY 586 ACTGATGTCACTGGCTCAATAGCTTTAGCAATTTGGATTTCTTGTTCGAATTCGACATTTA 645  
Db 161 ThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGlyHisLeu 180  
QY 646 TTTGCAATCAATATATCTGGTGCACGATGAATCCCGCCGATTCCTTTGGACCTGCAGTT 705  
Db 181 PheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyProAlaVal 200  
QY 706 ATCATGGGAAATGGGAAACCATTTGGATATATTTGGTGGCCCATCATAGAGCTGTC 765  
Db 201 IleMetGlyAsnTrpAlaAsnHisIleIleTyrTrpValGlyProIleMetGlyAlaVal 220  
QY 766 CTCGCTGGTGGCTTTATGATGATGTCTTCTGTCCAGATGTTGAATTCAAAGCTGCTTT 825  
Db 221 LeuAlaGlyAlaLeuTyrGlnTyrValPheCysProAspValGluLeuLysArgLeu 240  
QY 826 AAGAAGCTTCCAGCAAACTGCCAGCAAAACAAAGAAAGCTTACATGAGGTGGAGGAC 885  
Db 241 LysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluValGluAsp 260  
QY 886 AACAGAGTCAGTAGAGCGATGACCTGATTTCTTAAACCTGGAGTGTGTCATGATT 945  
Db 261 AsnArgSerGlnValGluThrGluAspLeuIleLeuLysProGlyValValHisValIle 280  
QY 946 GACGTTGACCGGGAGAGAGAGAAAGCGGAAAGCAATCTGGAGAGGTATTGTTCTCA 1005  
Db 281 AspIleAspArgGlyGluGluLysLysGlyLysAspSerSerGlyGluValLeuSerSer 300  
QY 1006 GTA 1008  
Db 301 Val 301

## RESULT 12

US-10-732-923-15580  
; Sequence 15580, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; PRIOR FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 15580  
; LENGTH: 300  
; TYPE: PRF  
; ORGANISM: Mus musculus  
US-10-732-923-15580

## Alignment Scores:

Pred. No.: 1,81e-132 Length: 300  
Score: 1458.50 Matches: 281  
Percent Similarity: 96.01% Conservative: 8  
Best Local Similarity: 93.36% Mismatches: 11  
Query Match: 66.87% Indels: 1  
DB: 17 Gaps: 1

US-10-732-180-1 (1-1152) x US-10-732-923-15580 (1-300)

QY 106 ATGGTGGCTTCAAGGGCTCTGGACTCAAGCTTCTGGAAGAGCTACAGGGAATTT 165  
Db 1 MetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValSerAlaGluPhe 20  
QY 166 CTGGCCATGCTTATTTTCTCTCCTCAGCTGGGATCCACCATCAACTCACTGGGCTGAACA 225

Db 21 LeuAlaThrLeuIlePheVal---LeuGlyValGlySerThrIleAsnTrpGlyGlySer 39  
QY 226 GAAAAGCCTTTACCGGTGCACATGTTCTCTCATCTCCCTTGTCTTGGAGCTCAGCATGCA 285  
Db 40 GluAsnProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAla 59  
QY 286 ACCATGGTGCAGTGTCTTTGGCCATATCAGCGTGGCCACATCAACCCCTGCAGTGACTGTG 345  
Db 60 ThrMetValGlnCysLeuGlyHisIleSerGlyGlyHisIleAsnProAlaValThrVal 79  
QY 346 GCCATGGTGTGCACAGGAAGATCAGCATCCCAAGTCTGTCTTCTATCATCCAGCCCCAG 405  
Db 80 AlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIleIleAlaGln 99  
QY 406 TCCCTGGGGGCCCATCATTTGGAGCAGGAATCTCTATCTGGTCACACCTCCCGAGTGTGTG 465  
Db 100 CysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValThrProSerValVal 119  
QY 466 GGAGCCCTGGGAGTCACCATGGTTCATGGAATCTTACCGCTGGTGCATGGTCTCCTGGTT 525  
Db 120 GlyGlyLeuGlyValThrValHisGlyAsnLeuThrAlaGlyHisGlyLeuLeuVal 139  
QY 526 GAGTTGATTAATCACATTTCAATTTGGTGTCTTACTATCTTTGCCAGCTGTGATTTCCAAACGG 585  
Db 140 GluLeuIleIleThrPheGlnLeuValPheThrValPheAlaSerCysAspSerLysArg 159  
QY 586 ACTGATGTCACTGGCTCAATAGCTTTAGCAATTTGGATTTCTTGTTCGAATTCGACATTTA 645  
Db 160 ThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGlyHisLeu 179  
QY 646 TTTGCAATCAATTAATCTACTGGTGCAGCATGAATCCCGCCGATCTTTGGACCTGCAGTT 705  
Db 180 PheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyProAlaVal 199  
QY 706 ATCATGGGAAATTTGGGAAACCATTTGGATATATTTGGGTTGGGCCCATCATAGGAGCTGTC 765  
Db 200 IleMetGlyAsnTrpAlaAsnHisIleTyrTrpValGlyProIleMetGlyAlaVal 219  
QY 766 CTCGCTGGTGGCTTTATGATGATGTCTTCTGTCCAGATGTTGAATTCAAAGCTGCTTT 825  
Db 220 LeuAlaGlyAlaLeuTyrGlnTyrValPheCysProAspValGluLeuLysArgLeu 239  
QY 826 AAGAAGCTTTCAGCAAACTGCCAGCAAAACAAAGAAAGCTACATGAGGTGGAGGAC 885  
Db 240 LysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluValGluAsp 259  
QY 886 AACAGAGTCAGGTAGACGCGATGACCTGATTTCTTAAACCTGGAGTGTGTCATGATT 945  
Db 260 AsnArgSerGlnValGluThrGluAspLeuIleLeuLysProGlyValValHisValIle 279  
QY 946 GACGTTGACCGGGAGAGAGAGAAAGCGGAAAGCAATCTGGAGAGGTATTGTTCTCA 1005  
Db 280 AspIleAspArgGlyGluGluLysLysGlyLysAspSerSerGlyGluValLeuSerSer 299  
QY 1006 GTA 1008  
Db 300 Val 300

## RESULT 13

US-10-732-923-15746  
; Sequence 15746, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 15746

```

; LENGTH: 249
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-10-732-923-15746

Alignment Scores:
Pred. No.: 1-57e-54 Length: 249
Score: 656.50 Matches: 118
Percent Similarity: 69.72% Conservative: 34
Best Local Similarity: 54.13% Mismatches: 63
Query Match: 30.10% Indels: 3
DB: 17 Gaps: 1

```

US-10-723-180-1 (1-1152) x US-10-732-923-15746 (1-249)

```

QY 142 TGGAAAGCAGTCACAGCGGAATTTCTGGCCATCTTATTTTGTCTCTCAGCGTGGGA 201
Db 23 TtpArgMetLeuValAlaGluPheLeuGlyThrPheLeuValSerileGlyIleGly 42
QY 202 TCCACCATCACTGGGGTGGACAGAAAGCCTTTACCGTGCACATGTTCTCATCTCC 261
Db 43 SerThrMetGlyTrpGlyGlyAspTyrAlaPro-----ThrMetThrGlnileAla 59
QY 262 CTTTGTCTTGGAGTCAGCATTCGACCATGGTGCAGTGTGGCCATATCAGCGGTGC 321
Db 60 PheThrPheGlyLeuValAlaThrLeuAlaGlnAlaPheGlyHisValSerileGlyCys 79
QY 322 CACATCAACCCCTGCAGTGTGGCCATGGTGTGCACAGGAAGATCAGCATGCCAAG 381
Db 80 HisIleAsnProAlaValThrIleGlyLeuMetIleThrAlaAspIleSerileLeuLys 99
QY 382 TCTGTCTTACATCGCAGCCAGTGCCTGGGGCCATCATTTGGAGCAGGAATCTCTAT 441
Db 100 GlyAlaPheTyrIleValSerGlnCysValGlyAlaIleAlaGlyAlaAlaLeuLys 119
QY 442 CTGGTCAACACCTCCAGTGTGGTGGAGCCCTGGAGTCACCATGGTTCATGGAATCTT 501
Db 120 AlaAlaThrProSerAspValIleGlyLeuGlyValThrGlyIleAspProargLeu 139
QY 502 ACCGTGTGTATGCTCTCTGTGTGATGATGATGATGATGATGATGATGATGATGATG 561
Db 140 ThrAlaGlyGlnGlyValMetMetGluAlaLeuIleThrPheIleLeuValPheVal 159
QY 562 TTTGCCAGTGTGATTCACAAAGGAGTGTGATGATGATGATGATGATGATGATGATG 621
Db 160 HisGlyValCysAspAsnArgSerAspIleLysGlySerAlaProLeuAlaIleGly 179
QY 622 TTTTCTGTGCAATTTGACATTTATTTGCAATTTATTTGCAATTTATTTGCAATTT 681
Db 180 LeuSerIleThrAlaGlyHisLeuSerAlaIleLysTyrThrGlyAlaSerMetAsnPro 199
QY 682 GCCGATCTTTGGACCTGACATGATGATGATGATGATGATGATGATGATGATGATG 741
Db 200 AlaArgSerPheGlyProAlaValValMetGlyAsnTyrThrAspGlnTyrValTyr 219
QY 742 GTTGGGCCCATCATAGGAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTC 795
Db 220 ValGlyProIleValGlyGlyIleValGlyAlaValTyrArgLeuPhePhe 237

```

RESULT 14

```

US-10-732-923-15615
; Sequence 15615, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2003-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15615

```

```

; LENGTH: 251
; TYPE: PRT
; ORGANISM: Anopheles gambiae str. PEST
US-10-732-923-15615

```

```

Alignment Scores:
Pred. No.: 1-82e-51 Length: 251
Score: 625.00 Matches: 114
Percent Similarity: 67.43% Conservative: 33
Best Local Similarity: 52.29% Mismatches: 67
Query Match: 28.66% Indels: 4
DB: 17 Gaps: 1

```

US-10-723-180-1 (1-1152) x US-10-732-923-15615 (1-251)

```

QY 142 TGGAAAGCAGTCACAGCGGAATTTCTGGCCATCTTATTTTGTCTCTCAGCGTGGGA 201
Db 24 TtpArgMetLeuValAlaGluPheLeuGlyThrPheLeuValAlaIleGlyIleGly 43
QY 202 TCCACCATCACTGGGGTGGAAACAGAAAGCCTTTACCGTGCACATGTTCTCATCTCC 261
Db 44 SerThrThrGlyTyrThrAspTyrSerProThrLeu-----ThrGlnileAla 59
QY 262 CTTTGTCTTGGAGTCAGCATTCGACCATGGTGCAGTGTGGCCATATCAGCGGTGC 321
Db 60 PheThrPheGlyLeuValAlaThrLeuAlaGlnAlaPheGlyHisValSerileGlyCys 79
QY 322 CACATCAACCCCTGCAGTGTGGCCATGGTGTGCACAGGAAGATCAGCATGCCAAG 381
Db 80 HisIleAsnProAlaValThrIleGlyLeuIleValThrAlaAspValSerileLeuLys 99
QY 382 TCTGTCTTACATCGCAGCCAGTGCCTGGGGCCATCATTTGGAGCAGGAATCTCTAT 441
Db 100 GlyAlaPheTyrIleValSerGlnCysIleGlyAlaIleAlaGlyAlaAlaValIleLys 119
QY 442 CTGGTCAACACCTCCAGTGTGGTGGAGCCCTGGAGTCACCATGGTTCATGGAATCTT 501
Db 120 AlaAlaThrProSerGluValValGlyLeuGlyValThrGlyIleAlaProGlyLeu 139
QY 502 ACCGTGTGTATGCTCTCTGTGTGATGATGATGATGATGATGATGATGATGATGATG 561
Db 140 SerThrGlyGlnGlyValLeuIleGluAlaLeuIleThrPheMetLeuValPheVal 159
QY 562 TTTGCCAGTGTGATTCACAAAGGAGTGTGATGATGATGATGATGATGATGATGATG 621
Db 160 HisGlyValCysAspAsnArgSerAspValLysGlySerAlaProLeuAlaIleGly 179
QY 622 TTTTCTGTGCAATTTGACATTTATTTGCAATTTATTTGCAATTTATTTGCAATTT 681
Db 180 LeuSerIleThrAlaGlyHisLeuAlaIleLysTyrThrGlyAlaSerMetAsnPro 199
QY 682 GCCGATCTTTGGACCTGACATGATGATGATGATGATGATGATGATGATGATGATG 741
Db 200 AlaArgSerPheGlyProAlaValValMetGlyAsnTyrThrAspLeuTyrValTyr 219
QY 742 GTTGGGCCCATCATAGGAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTC 795
Db 220 ValGlyProIleValGlyGlyIleValGlyAlaValTyrArgLeuPhePhe 237

```

RESULT 15

```

US-10-097-340-12
; Sequence 12, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN

```

```
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-12
```

```
Alignment Scores:
Pred. No.: 1,11e-50 Length: 265
Score: 617.00 Matches: 132
Percent Similarity: 64.64% Conservative: 38
Best Local Similarity: 50.19% Mismatches: 81
Query Match: 28.29% Indels: 12
DB: 14 Gaps: 4
```

US-10-723-180-1 (1-1152) x US-10-097-340-12 (1-265)

```
QY 118 AAAGGGGTCTGAGCTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTT 177
||| ||| :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3 LysGluValCysSerValAlaPheLeuLysAlaValPheAlaGluPheLeuAlaThrLeu 22

QY 178 ATTTTGTCTCTCCTCAGCCTGGGATCCACATCAACTGGGGTGGAAACAGAAAAGCCTTTA 237
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23 IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp-----ProSerAlaLeu 39

QY 238 CCGGTCGACATGTTCTCATCTCCCTTGTGGACTCAGCATTCACACCATGGTGCGAG 297
||| :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 40 ProThr---IleLeuGlnIleAlaLeuAlaPheGlyLeuAlaIleGlyThrLeuAlaGln 58

QY 298 TGCTTTGGCCATATCAGCGTGGCCACATCAACCTGCAGTGACTGTGGCCATGGTGTC 357
||| :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59 AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuVal 78

QY 358 ACCAGGAAGATCAGCATCGCCAAAGTCTGTCTTCTACATCGCAGCCAGTCCTGGGGGCC 417
||| :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 79 GlyAsnGlnIleSerLeuLeuArgAlaPhePheTy-ValAlaAlaGlnLeuValGlyAla 98

QY 418 ATCATTTGGAGCAGGAATCTCTATCTGTCTGGTCACACCTCCAGTGTGGGAGCCCTGGGA 477
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 99 IleAlaGlyAlaGlyIleLeuTyGlyValAlaProLeuAsnAlaAlaArgGlyAsnLeuAla 118

QY 478 GTCAACATGTTTCATCGAATCTTACCGCTGGTGCTCCTCGTGTGATTAATC 537
||| :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 119 ValAsnAlaLeuAsnAsnThrThrGlnGlyGlnAlaMetValValGluLeuIleLeu 138

QY 538 ACATTTCAATTGGTGTACTATCTTTTGGCAGCTGTGATTCCAAACGGACTGTGTCACT 597
```

```
Db 139 ThrPheGlnLeuAlaLeuCysIlePheAlaSerThrAspSerArgThrSerProVal 158
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 598 GGTCAATAGCTTTAGCAATTGGATTCTGTGTGCAATTGGACATTTATTGCAATCAAT 657
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTy 178
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 658 TATACTGGTGCCAGCATGAATCCCGCCGATCCTTTGGACCTGCAGTTATCATGGGAAT 717
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValValMetAsnArg 198
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 718 TCGGAA---AACCATTTGATATATTGGTTGGGCCCATCATAGGAGCTGTCTCGTGGT 774
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 199 PheSerProAlaHisTrpValPheTrpValGlyProIleValGlyAlaValLeuAla 218
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 775 GGCCCTTTATGAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTAAAGAAGCC 834
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 219 IleLeuTyPheTyTrpLeuLeuPheProAsn-----SerLeuSer 231
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 835 TTCAGCAAGCTGCCCGAGCAAAACAAAAGGAAGCTACATGGAGGTGGAGGACCAACAGGAGT 894
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 232 LeuSerGluArgValAlaIleIleLysGlyThrTyGluProAspGluAspTrpGluGlu 251
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 895 CAGGTAGAG 903
||| |||||
Db 252 GlnArgGlu 254

Search completed: September 8, 2005, 12:23:27
Job time : 200 secs
```

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 8, 2005, 11:55:58 ; Search time 50.5 Seconds  
(without alignments)  
4389.771 Million cell updates/sec

Title: US-10-723-180-1  
Perfect score: 2181  
Sequence: 1 9999caggcaatgagagctg.....attactcagctctaaacaata 1152

Scoring table: BLOSOM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DBV=xlp  
-Q=/cpn2 1/USPTO spool\_p/US10723180/runat 08092005 123256 29250/app\_query.fasta\_1.1351  
-DB=PIR -QPMF=afastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10723180 @CNC 1 1 64 @runat 08092005 123256 29250 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DSV TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR.79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	1689.5	77.5	341	2 I39178	aquaporin 4, long
2	1595	73.1	323	2 I59283	water channel prot
3	614	28.2	265	2 A55630	aquaporin-5, saliv
4	612	28.1	263	2 A55279	major intrinsic pr
5	606	27.8	263	1 MMBOLM	lens fiber membran
6	602	27.6	262	2 JN0557	lens fiber membran
7	600	27.5	261	2 S53427	major intrinsic pr
8	597	27.4	271	2 A53442	aquaporin 2 - huma
9	594	27.2	271	2 J07050	water channel prot
10	589.5	27.0	272	2 I51164	chip aquaporin - e
11	589	27.0	271	2 I64818	water-channel aqua
12	573	26.3	269	2 B44499	major intrinsic pr
13	572	26.2	269	2 JC1320	water channel prot
14	569	26.1	271	2 JC2348	water channel prot

15	563	25.8	269	2 I52366	uterine water chan
16	561	25.7	269	2 A41616	erythrocyte integr
17	443.5	20.3	278	2 A84545	hypothetical prote
18	443.5	20.3	309	2 T14889	membrane intrinsic
19	439.5	20.2	286	2 T14601	plasma membrane ma
20	437	20.0	285	2 T12435	probable plasma me
21	436	20.0	286	2 T02451	plasma membrane in
22	434	19.9	286	2 T47935	plasma membrane in
23	434	19.9	290	2 T02879	probable plasma me
24	433.5	19.9	283	2 T12434	probable plasma me
25	433	19.9	286	2 A86147	hypothetical prote
26	430.5	19.7	281	2 T14600	plasma membrane ma
27	430	19.7	295	2 T01528	probable plasma me
28	429.5	19.7	281	2 T09124	probable aquaporin
29	429.5	19.7	287	2 T12440	mipC protein - com
30	427.5	19.6	280	2 T05780	plasma membrane in
31	425.5	19.5	248	2 T12632	water channel prot
32	425	19.5	284	2 T12557	mipE protein - com
33	425	19.5	286	2 S42542	ripening-associate
34	424	19.4	288	2 T09794	major intrinsic pr
35	418.5	19.2	250	2 T14000	aquaporin TIP7 - c
36	418	19.2	287	2 T09791	drought-induced ma
37	418	19.2	700	2 S09699	bib protein - frui
38	415.5	19.1	288	2 T14863	porin mipl - Norwa
39	415	19.0	288	2 S41194	transmembrane prot
40	414.5	19.0	248	2 T14001	aquaporin TIP18 -
41	413	18.9	286	2 T04367	plasma membrane in
42	412.5	18.9	248	2 T10804	tonoplast intrinsi
43	412.5	18.9	288	2 T14599	plasma membrane ma
44	412	18.9	287	2 T05378	probable plasma me
45	410.5	18.8	285	2 E84789	hypothetical prote

ALIGNMENTS

RESULT 1

I39178  
aquaporin 4, long splice form - human  
N:Alternate names: mercurial-insensitive water channel protein; MWC protein  
N:Contains: aquaporin 4, long splice form; aquaporin 4, medium splice form; aquaporin 4,  
N:Species: Homo sapiens (man)  
C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 09-Jul-2004  
C:Accession: I39178; I39177; JG6162; S66273  
R:Yang, B.; Ma, T.; Verkman, A.S.  
J. Biol. Chem. 270, 22907-22913, 1995  
A:Title: cDNA cloning, gene organization, and chromosomal localization of a human mercur  
A:Reference number: I39177; MUID:96032721; PMID:7559426  
A:Accession: I39178  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-341 <YANI>  
A:Cross-references: UNIPROT:P55087; EMBL:U34846; NID:G1072054; PIDN:AAC52112.1; PID:G107  
A>Note: alternatively spliced long form with upstream start codon  
A:Accession: I39177  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 41-341 <YAN2>  
A:Cross-references: EMBL:U34845; NID:G1072052; PIDN:AAC50284.1; PID:G1072053  
R:Lu, M.; Lee, M.D.; Smith, B.L.; Jung, J.S.; Agre, P.; Verdijk, M.A.J.; Merx, G.; Rijs  
Proc. Natl. Acad. Sci. U.S.A. 93, 10908-10912, 1996  
A:Title: The human AQP4 gene: Definition of the locus encoding two water channel polypep  
A:Reference number: JG6162; MUID:97008105; PMID:8855281  
A:Accession: JG6162  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 19-263, 'G', 265-304, 'VE', 307-313, 'P', 315-341 <LUA>  
A:Cross-references: GB:U63622; NID:G1680707; PIDN:AB26957.1; PID:G1680708  
R:Misaka, T.; Abe, K.; Iwabuchi, K.; Kusakabe, Y.; Ichinose, M.; Miki, K.; Emori, Y.; A  
FEBS Lett. 381, 208-212, 1996  
A:Title: A water channel closely related to rat brain aquaporin 4 is expressed in acid-  
A:Reference number: S66273; MUID:96176324; PMID:8601457  
A:Accession: S66273  
A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA  
 A:Residues: 19-263, 'G', 265-304, 'VE', 307-313, 'P', 315-341 <MIS>  
 A:Cross-references: GB:D63412; NID:g1236245; PIDN:BA09715.1; PID:g1236246  
 C:Comment: This protein plays a role in the pathophysiology of normal pressure hydrocephalus  
 C:Genetics:

A:Gene: GDB:AQP4; MIM:600308  
 A:Cross-references: GDB:371722; OMIM:600308  
 A:Map position: 18q11.2-18q12.1  
 A:Introns: 29/2; 167/3; 222/3; 249/3  
 C:Superfamily: lens fiber membrane major intrinsic protein  
 C:Keywords: alternative initiators; alternative splicing; glycoprotein; transmembrane protein  
 F1-341/Product: aquaporin 4, long splice form #status predicted <M1>  
 F19-341/Product: aquaporin 4, medium splice form #status predicted <MAT2>  
 F41-341/Product: aquaporin 4, short splice form #status predicted <MAT3>  
 F56-71/Domain: transmembrane #status predicted <TM1>  
 F98-111/Domain: transmembrane #status predicted <TM2>  
 F115-117/Region: NPA motif  
 F133-154/Domain: transmembrane #status predicted <TM3>  
 F176-195/Domain: transmembrane #status predicted <TM4>  
 F204-224/Domain: transmembrane #status predicted <TM5>  
 F231-233/Region: NPA motif  
 F250-270/Domain: transmembrane #status predicted <TM6>  
 F171,224,301/Binding site: carbohydrate (Asn) (covalent) #status predicted

## Alignment Scores:

Pred. No.:	6,296-133	Length:	341
Score:	1689.50	Matches:	326
Percent Similarity:	97.04%	Conservative:	2
Best Local Similarity:	96.45%	Mismatches:	5
Query Match:	77.46%	Indels:	5
DB:	2	Gaps:	1

US-10-723-180-1 (1-1152) x I39178 (1-341)

QY	10	AATGAGAGCTGC-----	ACTCTGGCTGGGAAAGGATGATGACAGACC	54
Db	4	AsnHisAlaCysPheValGluThrProAsnLeuAlaGlyGluGlyMetSerAspArgPro		23
QY	55	ACAGCAAGCGGTGGGTGAAGTGTGGACCTTTGTGTGACAGAGAACATCATGTGGCT		114
Db	24	ThrAlaArgArgTrpGlyLysCysGlyProLeuCysThrArgGluAsnIleMetValAla		43
QY	115	TTCAAGAGGGTCTGACTCAAGCTTCTGAAAGGAGTGCACAGCGGAATTTCTGCCCATG		174
Db	44	PheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAlaGluPheLeuAlaMet		63
QY	175	CTTATTTTGTCTCTCAGCTGGATCCACATCACTGGGTGGTGAACAGAAAGCCT		234
Db	64	LeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGlyThrGluLysPro		83
QY	235	TTACGGTTCGACATGTTCTCATCTCCCTTTGCTTTGGACTCAGCATTCACACCATGGTG		294
Db	84	LeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAlaThrMetVal		103
QY	295	CAGTCTTTGGCCATATCAGCGGTGGCCACATCAACCTTCGAGTGAATCTGGCCATGGTG		354
Db	104	GlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaValThrValAlaMetVal		123
QY	355	TGCACAGAGATCAGCATCCCAAGTCTGCTTCTACATCCGACCCAGTCCCTGGGG		414
Db	124	CysThrArgLysIleSerIleAlaLysSerValPheTyrlleAlaAlaGlnCysLeuGly		143
QY	415	GCATCATTTGGAGAGGAATCTCTATCTGTGTGCACACTCCAGTGTGTGGAGCCCTG		474
Db	144	AlaIleIleGlyAlaGlyIleLeuTyrlleValThrProProSerValValGlyGlyLeu		163
QY	475	GGAGTCACCATGTTTCATCGAATCTTACCGTGTGTCTGCTGCTGCTGCTGCTGATA		534
Db	164	GlyValThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeuLeuValGluLeuIle		183
QY	535	ATCACATTTCAATGGTGTATTAATCTTTTGGCAGCTGTGATTCACAAAGGACTGATGC		594
Db	184	IleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSerLysArgThrAspVal		203

QY	595	ACTGCTCAATAGACTTTAGCAATTGGATTTCTGTGTGCAATTGGACATTTATTTCATATC	654
Db	204	ThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGlyHisLeuPheAlaIle	223
QY	655	AATTATACCTGGTCCGAGCATGAATCCCGCCGATCTTTGGACCTGGAGTTATCATGGGA	714
Db	224	AsnTyrlThrGlyAlaSerMetAsnProAlaArgSerPheGlyProAlaValIleMetGly	243
QY	715	AATTCGGAAAAACCATTTGATATATTGGGTGGGCCCATCATAGGAGCTGTCTCTCGCTGT	774
Db	244	AsnTrpGluAsnHisTrpIleTyrlTrpValGlyProIleIleGlyAlaValLeuAlaGly	263
QY	775	GGCCTTTATGATGTCTTCTGTCACAGATGTGAATTCAAACGTCGTTTAAAGAACGCC	834
Db	264	AlaLeuTyrlGluTyrlValPheCysProAspValGluPheLysArgArgPheLysGluAla	283
QY	835	TTACGCAAGCTGCCAGCAACAAACAAAGAGCTACATGGAGGTGGAGCACACAGGAGT	894
Db	284	PheSerLysAlaAlaGlnGlnThrLysGlySerTyrlMetGluValGluAspAsnArgSer	303
QY	895	CAGGTAGAGACGAGTACCTGATTCATAAACTGGAGTGGTGCATGTGATTCAGCTGCAC	954
Db	304	GlnAlaLysThrAspAspLeuLeuLysLeuGlyValValHisValIleAspValAsp	323
QY	955	CGGGAG	1008
Db	324	ArgGlyGluGluLysLysGlyLysAspGlnSerGlyGluValLeuSerSerVal	341

## RESULT 2

159283

water channel protein, mercurial-insensitive - rat

N:Alternate names: aquaporin 4; MIMC

C:Species: Rattus norvegicus (Norway rat)

C&gt;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004

C:Accession: I59283; A53194

J:Jung, J.S.; Bhat, R.V.; Preston, G.M.; Guggino, W.B.; Baraban, J.M.; Agre, P.

Proc. Natl. Acad. Sci. U.S.A. 91, 13052-13056, 1994

A:Title: Molecular characterization of an aquaporin cDNA from brain: candidate osmorecep

A:Reference number: I59283; MUID:95108097; PMID:7528931

A:Accession: I59283

A&gt;Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-323 &lt;RES&gt;

A:Cross-references: UNIPROT:P47863; EMBL:U14007; NID:g595787; PIDN:AAC52152.1; PID:g5957

A:Experimental source: brain

R:Hasegawa, H.; Ma, T.; Skach, W.; Matthay, M.A.; Verkman, A.S.

J. Biol. Chem. 269, 5497-5500, 1994

A:Title: Molecular cloning of a mercurial-insensitive water channel expressed in selecte

A:Reference number: A53194; MUID:94164885; PMID:7509789

A:Accession: A53194

A:Molecule type: mRNA

A:Residues: 23-200, 'P', 202-323 &lt;HAS&gt;

A:Cross-references: GB:L27588; NID:g459950; PIDN:AAA17730.1; PID:g459951

C:Genetics:

A:Gene: AQP4

C:Superfamily: lens fiber membrane major intrinsic protein

C:Keywords: glycoprotein; phosphoprotein; transmembrane protein

F:30-57/Domain: transmembrane #status predicted &lt;TM1&gt;

F:70-89/Domain: transmembrane #status predicted &lt;TM2&gt;

F:115-136/Domain: transmembrane #status predicted &lt;TM3&gt;

F:159-177/Domain: transmembrane #status predicted &lt;TM4&gt;

F:185-205/Domain: transmembrane #status predicted &lt;TM5&gt;

F:231-252/Domain: transmembrane #status predicted &lt;TM6&gt;

F:107/Binding site: phosphate (Thr) (covalent) #status predicted

F:153/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:180/Binding site: phosphate (Ser) (covalent) #status predicted

## Alignment Scores:

Pred. No.:	4,84e-125	Length:	323
Score:	1595.00	Matches:	304
Percent Similarity:	96.59%	Conservative:	8
Best Local Similarity:	94.12%	Mismatches:	11



```
Query Match: 73.13% Indels: 0
DB: 2 Gaps: 0
US-10-723-180-1 (1-1152) x I59283 (1-323)
QY 40 ATGAGTGACAGACCCACAGACCGGTGGGTAAAGTGGACCTTTGTGTACCAGAG 99
DB 1 MetSerAspGlyAlaAlaAlaArgArgTrpGlyLysCysGlyProCysSerArgGlu 20
QY 100 AACATCATGTGGCTTTCAAAGGGGTCTGAGCTCAAGCTTTCTGGAAGCAGTCAACGCG 159
DB SerIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAla 40
QY 160 GAATTTCTGGCCATGCTTATTTTGTCTCTCAGCTCGGAGTCCACCATCAACTGGGT 219
DB GluPheLeuAlaMetLeuIlePheValLeuLeuSerValGlySerThrIleAsnTrpGly 60
QY 220 GGAACAGAAAGCCTTTACGGTCGACATGGTTCTCATCTCCCTTTGCTTTGACATCAGC 279
DB GlySerGluAsnProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer 80
QY 280 ATTGCAACCATGTGTCAGTGTCTTTGGCCATATCAGCGGTGGCCACATCAACCTCGAGTG 339
DB IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100
QY 340 ACTGTGGCCATGGTGTGCCAGGAAAGATCAGCATCGCCCAAGTCTGTCTTCTACATCGCA 399
DB ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIleThr 120
QY 400 GCCAGTGTCTGGGGCCCATCATTTGAGCAGGAGNATCTCTATCTGTGTCACTCCCACT 459
DB AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValThrProSer 140
QY 460 GTGGTGGGAGCCTGGAGTCCACATGGTTTCATGGAAATCTTACCGCTGTATGTCTC 519
DB ValValGlyGlyLeuGlyValThrValHisGlyAsnLeuThrAlaGlyHisGlyLeu 160
QY 520 CTGGTTGAGTTGATAATCAATTTCAATTTGGTGTCTTACTATCTTTGCCAGCTGTGATCC 579
DB LeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180
QY 580 AACCGACTGATGTCTACTGGCTCAATAGCTTTAGCAATGGATTTCTGTGGCAATTGGA 639
DB LysArgThrAspValThrGlySerValAlaLeuAlaIleGlyPheSerValAlaIleGly 200
QY 640 CATTTATTTGCAATCAATATATCTGTGTGCAGCATGAATCCCGCCGATCTTGGACCT 699
DB HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220
QY 700 GCAGTTTATCATGGAAATTTGGGAAACCATTTGGATATATTTGGTGGGCCCATCATAGGA 759
DB AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly 240
QY 760 GCTGTCTCGTGTGGCTTTATAGTAGTAGTCTTCTGTCCAGATGTGAATTCAAACCT 819
DB AlaValLeuAlaGlyAlaLeuTyrGluTyrValPheCysProAspValGluLeuLysArg 260
QY 820 CTTTTTAAAGAGCCTTCAGCAAGCTGCCAGCAACCAAAAGGAGCTACATGGAGTG 879
DB ArgLeuLysGluAlaPheSerLysAlaAlaGlnThrLysGlySerTyrMetGluVal 280
QY 880 GAGGACAAACGAGTTCAGTAGACGGATGACCTGATTTCTAAACCTGGAGTGGTGCAT 939
DB GluAspAsnArgSerGlnValGluThrGluAspLeuIleLeuLysProGlyValValHis 300
QY 940 GTGATTGACGTTGACCGGGAGAGAGAAAGAGGAAAGACCAATCTCGAGAGGTATTG 999
DB ValIleAspIleAspArgGlyAspGluLysLysGlyLysAspSerSerGlyGluValLeu 320
QY 1000 TCTTCAGTA 1008
DB SerSerVal 323
```

## RESULT 3

```
A55630
aquaporin-5, salivary gland - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 09-Jul-2004
C:Accession: A55630
R:Raina, S.; Preston, G.M.; Guggino, W.B.; Agre, P.
J. Biol. Chem. 270, 1908-1912, 1995
A:Title: Molecular cloning and characterization of an Aquaporin cDNA from salivary, lac-
A:Reference number: A55630; MUID:95130575; PMID:7530250
A:Accession: A55630
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-265 <RAI>
A:Cross-references: UNIPROT:P47864; GB:U16245; NID:G664759; PIDN:AAA66221.1; PID:G664759
C:Superfamily: lens fiber membrane major intrinsic protein
C:Keywords: transmembrane protein

Alignment Scores:
Pred. No.: 3,72e-43 Length: 265
Score: 614.00 Matches: 133
Percent Similarity: 56.83% Conservative: 45
Best Local Similarity: 42.22% Mismatches: 82
Query Match: 28.15% Indels: 54
DB: 2 Gaps: 5

US-10-723-180-1 (1-1152) x A55630 (1-265)
QY 118 AAAGGGTCTGCACTCAAGCTTTCTGGAAAGCAGTCAACGGGAATTTCTGCCATGCTT 177
DB 3 LysGluValCysSerLeuAlaPhePheLysAlaValPheAlaGluPheLeuAlaThrLeu 22
QY 178 ATTTTGTCTCTCTCAGCTGGATCCACCATCACTGGGTGGACAGAAAGCCTTTA 237
DB 23 IlePheValPhePheGlyLeuGlySerAlaLeuLysTrp-----ProSerAlaLeu 39
QY 238 CGGTGCGACATGGTTCTCATCTCCCTTTGCTTTGGACTCAGCATTTGCAACCATGTGTCAG 297
DB 40 ProThr---IleLeuGlnIleSerIleAlaPheGlyLeuAlaIleGlyThrLeuAlaGln 58
QY 298 TGCTTTGGCCATATCAGCGGTGGCCACATCAACCTCGAGTCACTGTGGCCATGTGTGC 357
DB 59 AlaLeuGlyProValSerGlyGlyHisIleAsnProAlaIleThrLeuAlaLeuLeu 78
QY 358 ACCGGAAGATCAGCATCGCCAAAGTCTCTCTCATCTCGACCGACCGCTCTGGGGCC 417
DB 79 GlyAsnGlnIleSerLeuLeuArgAlaValPheTyrValAlaAlaGlnLeuValGlyAla 98
QY 418 ATCATTTGGACGAGGAATCTCTATCTGCTCACACCTCCAGTGTGTGGAGGCGCTGGGA 477
DB 99 IleAlaGlyAlaGlyIleLeuTyrTrpLeuAlaProLeuAsnAlaArgGlyAsnLeuAla 118
QY 478 GTCACATGTTTCATGGAAATCTTACCGCTGGTCAGGTCTCTCTGGTTGAGTTGATAATC 537
DB 119 ValAsnAlaLeuAsnAsnAsnThrThrProGlyLysAlaMetValValGluLeuIleLeu 138
QY 538 ACATTTCAATGGTGTCTTACTATCTTCCAGCTGTGATTCCAAACGGACATGATGTCAC 597
DB 139 ThrPheGlnLeuAlaLeuCysIlePheSerThrAspSerArgArgThrSerProVal 158
QY 598 GCCTCAATAGCTTTAGCAATTTGATTTCTGTTCATTTGGACATTTATTTTCAATCAAT 657
DB 159 GlySerProAlaLeuSerIleGlyLeuSerValThrLeuGlyHisLeuValGlyIleTyr 178
QY 658 TATACTGTGGCAGCATGAATCCCGCCGATCTCTTTGGACCTCGAGTTCATTCATGGAAAT 717
DB 179 PheThrGlyCysSerMetAsnProAlaArgSerPheGlyProAlaValValMetAsnArg 198
QY 718 TGGGAA---AACCATTTGGATATATTTGGTGGGCCCATCATAGGAGCTGTCTCTCGTGT 774
DB 199 PheSerProSerHisTrpValPheTrpValGlyProIleValGlyAlaMetLeuAla 218
QY 775 GGCCTTTATAGTATGTCTTCTGTCCAGATGTTGAATTCAAACGTCGTTTTTAAAGAGCC 834
```

```
Db 219 ILeuTyrPheTyrLeuLeuPheProSer-----SerLeuSer 231
      ||||| ||||| |||||
QY 835 TTCAGCAAAAGCTGCCAGCAAAACAAAGAAAGCTACATGAGAGTGGAGGACAAACAGGAGT 894
      ||||| ||||| ||||| ||||| |||||
Db 232 LeuHisAspArgValAlaValVallyysGlyThrTyrGluProGluGluAap----- 248
      ||||| ||||| ||||| ||||| |||||
QY 895 CAGGTAGACGATGACCTGATTCTAAACCTGGAGTGGTGCATGTGATTGACGTTGAC 954
      ||||| ||||| ||||| ||||| |||||
Db 248 ----- 248
QY 955 CGGGAGAGAGAGAAAGGAAAGACCAATCGGAGAGGTATTGTCTTCAGTATGACTA 1014
      ||||| ||||| ||||| ||||| |||||
Db 249 -----Tyr 249
QY 1015 GAAGATCGCACTGAAGACAGACAAAGACTCCTTAGAAGCTCTCTCA 1059
      ||||| ||||| ||||| ||||| |||||
Db 250 GluAspHisArgGluGluArgLysLysThrIleGluLeuThrAla 264
      ||||| ||||| ||||| ||||| |||||
RESULT 4
A55279
major intrinsic protein - human
N/Alternate names: MP26
C/Species: Homo sapiens (man)
C/Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C/Accession: A55279
R/Pisano, M.M.; Chappelinsky, A.B.
Genomics 11, 981-990, 1991
A/Title: Genomic cloning, complete nucleotide sequence, and structure of the human gene
A/Reference number: A55279; MUID:92147144; PMID:1840563
A/Accession: A55279
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-263 <PIS>
A/Cross-references: UNIPROT:P30301; GB:U36308; NID:G6579177; PIDN:AAC02794.2; PID:G65791
A/Experimental source: lens
A/Note: sequence extracted from NCBI backbone (NCBIN:80218, NCBI:P:80220)
C/Genetics:
A/Gene: GDB:MIP
A/Cross-references: GDB:119394; OMIM:154050
A/Map position: 12q13-12q13
C/Superfamily: lens fiber membrane major intrinsic protein
C/Keywords: eye lens; transmembrane protein
Alignment Scores:
Pred. No.: 5,46e-43 Length: 263
Score: 612.00 Matches: 114
Percent Similarity: 65.27% Conservative: 57
Best Local Similarity: 43.51% Mismatches: 77
Query Match: 28.06% Indels: 14
DB: 2 Gaps: 3
US-10-723-180-1 (1-1152) x A55279 (1-263)
QY 124 GTCCTG-----ACTCAAGCTTCTGGAAGCAGTCACAGCGGAATTCCTGGCCATG 174
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MetTrpLeuArgSerAlaSerPheTrpArgAlaIlePheAlaGluPheAlaThr 20
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 175 CTTATTATTGTTCTCTCAGCTGGATCCACATCAACTGGGCTGGGTGGAACAGAAAAGCCT 234
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 LeuPheTyrValPhePheGlyLeuGlySerSerLeuArgTrpAla-----Pro 36
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 235 TTACCGGTGCATGTTCTCATCTCCCTTTGAGTACATGCAATTCGACATTCGACATGGG 294
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37 GlyProLeuHisValLeuGlnValAlaMetAlaPheGlyLeuAlaLeuAlaThrLeuVal 56
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 295 CAGTGTCTTGGCCATATCAGCGTGGCCACATCAACCTGCAGTGCATGTGGCCATGNG 354
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 57 GlnSerValGlyHisIleSerGlyAlaHisValAsnProAlaValThrPheAlaPheLeu 76
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 355 TGCACGGAAGATCAGCATCGCCAAAGTCTGTCTTACATCGCAGCCGACGCTGGGG 414
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77 ValGlySerGlnMetSerLeuLeuArgAlaPheCysTyrMetAlaAlaGlnLeuGly 96
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 415 GCCATCTGGAGCAGGAATCCTCTATCTGGTCACACCTCCAGTGTGGTGGAGCGCTG 474
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97 AlaValAlaGlyAlaAlaValLeuTyrSerValThrProAlaValArgGlyAsnLeu 116
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 475 GGAGTCACCATGGTTTCATGGAAATCTTACCGCTCGTCATGGTCTCTCGTGTGAGTTGATA 534
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 117 AlaLeuAsnThrLeuHisProAlaValSerValGlyGlnAlaThrThrValGluIlePhe 136
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 535 ATCACTTCAATTGGTGTTPACTATCTTTGCCAGCTGTGATTCCTCAATTCGACATTTTGCATC 594
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137 LeuThrLeuGlnPheValLeuCysIlePheAlaThrTyrAspGluArgAsnGlyGln 156
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 595 ACTGCTCAATAGCTTTTAGCAATTTGATTTCTCTGCAATTCGACATTTTTCGATTCGATC 654
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 157 LeuGlySerValAlaLeuAlaValGlyPheSerLeuAlaLeuGlyHisLeuPheGlyMet 176
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 655 AATTATATCTGTCGCGCATCAATCCCGCCGATCTTTGGACCTGCGATTCATTCATGGGA 714
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 177 TyrThrThrGlyAlaGlyMetAsnProAlaArgSerPheAlaProAlaIleLeuThrGly 196
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 715 AATTGGGAAACCATGTGATATATTGGTTGGGCCCATCATAGAGCTGTCTCGCTGGT 774
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 197 AsnPheThrAsnHisTrpValTyrTrpValGlyProIleIleGlyGlyLeuGlySer 216
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 775 GGCCTTTATGAGTATGCTTCTCTGTCAGATGTTGAAATTCAAACCTGCTTTTAAAGAGGCC 834
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 217 LeuLeuTyrAspPheLeuLeuPheProArgLeu-----LysSer 229
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 835 TTCAGCAAGCTGCCAGCAAAACAAAGAGAGCTACATGGAGTGGAGGACAAACAGGAGT 894
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 230 IleSerGluArgLeuSerValLeuLysGlyAlaLysProAspValSerAsnGlyGlnPro 249
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 895 CAGGTA 900
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 250 GluVal 251
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 5
MMBOLM
lens fiber membrane major intrinsic protein - bovine
N/Alternate names: lens fiber membrane intrinsic 26K protein (MIP); MP26
C/Contains: MP22
C/Species: Bos primigenius taurus (cattle)
C/Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C/Accession: A23251; B23251; A22444
R/Gorin, M.B.; Yancey, S.B.; Cline, J.; Revel, J.P.; Horwitz, J.
Cell 39, 49-59, 1984
A/Title: The major intrinsic protein (MIP) of the bovine lens fiber membrane: characteri
A/Reference number: A23251; MUID:85024900; PMID:6207938
A/Accession: A23251
A/Molecule type: mRNA
A/Residues: 1-263 <GOR>
A/Cross-references: UNIPROT:P06624; GB:K02818; NID:G163296; PIDN:AAA30622.1; PID:G163297
A/Accession: B23251
A/Molecule type: protein
A/Residues: 1-39 <GO2>
R/Ingoc, L.D.; Faroutaud, P.; Dunia, I.; Benedetti, E.L.; Hoebeke, J.
FEBS Lett. 181, 74-78, 1985
A/Title: Sequence analysis of peptide fragments from the intrinsic membrane protein of c
A/Reference number: A22444; MUID:85127500; PMID:3882455
A/Accession: A22444
A/Molecule type: protein
A/Residues: 1-13, 'L',15-28, 'R',30, 'F',32, 'L',184-214 <NGO>
C/Comment: This protein is the major component of lens fiber gap junctions.
C/Superfamily: lens fiber membrane major intrinsic protein
C/Keywords: eye lens; gap junction; phosphoprotein; transmembrane protein
F/12-28/Domain: transmembrane #status predicted <TM1>
F/44-60/Domain: transmembrane #status predicted <TM2>
F/87-103/Domain: transmembrane #status predicted <TM3>
F/130-146/Domain: transmembrane #status predicted <TM4>
F/160-176/Domain: transmembrane #status predicted <TM5>
Alignment Scores:
```

Pred. No.: 1,73e-42 Length: 263  
 Score: 606.00 Matches: 113  
 Percent Similarity: 64.07% Conservative: 60  
 Best Local Similarity: 41.85% Mismatches: 83  
 Query Match: 27.79% Indels: 14  
 DB: 1 Gaps: 3

US-10-723-180-1 (1-1152) x MMBOLM (1-263)

```

QY 124 GTCGG-----ACTCAAGCTTTCTGGAAGCAGTCACAGCGGAATTTCTGGCCATG 174
Db 1 MetTrpGluLeuArgSerAlaSerPheTrpArgAlaIleCysAlaGluPheAlaSer 20

QY 175 CTTATTTTGTCTCTCAGCGTGGGATCCACCACTCAACTGGGGTGGGAACAGAAAGCCT 234
Db 21 LeuPheTyrValPhePheGlyLeuGlyAlaSerLeuArgTrpAla-----Pro 36

QY 235 TTACCGGTGCAGATGTTCTCATCTCCCTTGTGGTGGACTGACATGCAACCATGGTG 294
Db 37 GlyProLeuHisValLeuGlnValAlaLeuAlaPheGlyLeuAlaLeuAlaThrLeuVal 56

QY 295 CAGTGTCTTGGCCATATCAGCGTGGCCACATCAACCTGCAGTGTGGCCATGGTG 354
Db 57 GlnAlaValGlyHisIleSerGlyAlaHisValAsnProAlaValThrPheAlaPheLeu 76

QY 355 TGCACGGAAGATCAGCATCGCAAGTCTGTCTTACATCGCAGCCAGTCCCTGGGG 414
Db 77 ValGlySerGlnMetSerLeuLeuArgAlaIleCysTyMetValAlaGlnLeuLeuGly 96

QY 415 GCATCATTTGGAGCAGGAATCTCTATCTGTGTGCACACCTCCAGTGTGGGAGCCCTG 474
Db 97 AlaValAlaGlyAlaAlaValLeuTySerValThrProAlaValArgGlyAsnLeu 116

QY 475 GGAGTCACCATGTTTCATGGAATCTTACCGTGTGTCTCTGTGTGAGTTGATA 534
Db 117 AlaLeuAsnThrLeuHisProGlyValSerValGlyGlnAlaThrIleValGluIlePhe 136

QY 535 ATCACAATTTCAATGTGTGTTACTATCTTTTGGCAGCTGTGATTTCCAAACGGAGCTGATGTC 594
Db 137 LeuThrLeuGlnPheValLeuCysIlePheAlaThrTyArgGluArgAsnGlyArg 156

QY 595 ACTGGCTCAATAGCTTTAGCAATTTGATTTCTGTGCAATTTGAGCATTTATTTGCAATC 654
Db 157 LeuGlySerValAlaLeuAlaValGlyPheSerLeuThrLeuGlyHisLeuPheGlyMet 176

QY 655 AATTATACGTGGCCAGCATGAATCCCGCCCATCTTTGGACCTGACGTTATCATGGA 714
Db 177 TyrTyThrGlyAlaGlyMetAsnProAlaArgSerPheAlaProAlaIleLeuThrArg 196

QY 715 AATTGGAAACCATTTGGATATATTGGTTGGGCCCATCATAGGAGCTGTCTCGTGTG 774
Db 197 AsnPheThrAsnHisTrpValTyTrpValGlyProValIleGlyAlaGlyLeuGlySer 216

QY 775 GCGCTTTATGATGATGCTCTCTGTCAGAGTTTGAATTCAAACGTGTTTAAAGAGCC 834
Db 217 LeuLeuTyArgPheLeuLeuPheProArgLeu-----LysSer 229

QY 835 TTCAGCAAGCTGCCAGCAACAAAGAGGAGCTACATGAGGTGGAGGACACACAGAGT 894
Db 230 ValSerGluArgLeuSerIleLeuGlyGlySerArgProSerGluSerAsnGlyGlnPro 249

QY 895 CAGGTGAGAGCGGATGACCTGATTTCTAAAA 924
Db 250 GluValThrGlyGluProValGluLeuLeuLys 259

```

RESULT 6

JN0557  
 lens fiber membrane major intrinsic protein - African clawed frog  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 13-Aug-1999  
 C:Accession: JN0557  
 R:Austin, L.A.R.; Rice, S.J.; Baldo, G.J.; Lange, A.J.; Haspel, H.C.; Mathias, R.T.  
 Gene 124, 303-304, 1993

A:Title: The cDNA sequence encoding the major intrinsic protein of frog lens.  
 A:Reference number: JN0557; MUID:93185940; PMID:7916707  
 A:Accession: JN0557  
 A:Molecule type: mRNA  
 A:Residues: 1-262 <AUS>  
 A:Cross-references: EMBL:X56970; NID:g288514; PIDN:CAA0291.1; PID:g288515  
 A:Experimental source: lens  
 A:Superfamily: lens fiber membrane major intrinsic protein  
 C:Keywords: eye lens; gap junction; transmembrane protein

Alignment Scores:  
 Pred. No.: 3,73e-42 Length: 262  
 Score: 602.00 Matches: 116  
 Percent Similarity: 65.60% Conservative: 48  
 Best Local Similarity: 46.40% Mismatches: 77  
 Query Match: 27.60% Indels: 9  
 DB: 2 Gaps: 2

US-10-723-180-1 (1-1152) x JN0557 (1-262)

```

QY 136 GCTTTCTGGAAGCAGTCACAGCGGAATTTCTGGCCATGCTTATTTTGTCTCTCCTCAGC 195
Db 7 SerPheTrpArgAlaValPheAlaGluPhePheGlyThrMetPheTyValPhePheGly 26

QY 196 CTGGGATCCACCATCAACTGGGGTGGAAACAGAAAAGCCTTTACCGTCCGACATGGTTCTC 255
Db 27 LeuGlyAlaSerLeuLysTrpAlaAlaGly-----ProAlaAsnValLeuVal 42

QY 256 ATCTCCCTTTGCTTTGGACTCAGCATTCGCAACCATGGTGCAGTGTGGCCATATCAGC 315
Db 43 IleAlaLeuAlaPheGlyLeuValLeuAlaThrMetValGlnSerIleGlyHisValSer 62

QY 316 GTGGCCACATCAACCTGCAGTGTGGCCATGGTGTGCACACGAGGAGATCAGCATC 375
Db 63 GlyAlaHisIleAsnProAlaValThrPheAlaPheLeuIleGlySerGlnMetSerLeu 82

QY 376 GCCAAGTCTGTCTTACATCGCAGCCAGTGTCTGGGGCCCATCATTTGGAGCAGAAATC 435
Db 83 PheArgAlaIlePheTyIleAlaAlaGlnLeuLeuGlyAlaValAlaGlyAlaVal 102

QY 436 CTCATCTGTGTACACCTCCAGTGTGGTGGAGGCCCTGGGAGTCACCATGTTTATGGA 495
Db 103 LeuTyGlyValThrProAlaAlaIleArgGlyAsnLeuAlaLeuAsnThrLeuHisPro 122

QY 496 AATCTTACCGTGTGCATGGTCTCTGTTGAGTTGATAATCACATTTCAATTGTGTGTT 555
Db 123 GlyValSerLeuGlyGlnAlaThrThrValGluIlePheLeuThrLeuGlnPheValLeu 142

QY 556 ACTATCTTTGGCAGCTGTGATTTCCAAACGGACTGATGTCTCTGCTCAATAGCTTTAGCA 615
Db 143 CysIlePheAlaThrTyArgGluArgArgAsnGlyArgLeuGlySerValSerLeuAla 162

QY 616 ATTGATTTTCTGTGCAATTTGAGCATTTATTTGCAATTAATATATATGTTGTCAGCATG 675
Db 163 IleGlyPheSerLeuThrLeuGlyHisLeuPheGlyLeuTyTyThrGlyAlaSerMet 182

QY 676 AATCCCGCCCATCTTTGGACCTGCTATCATGGGAAATTTGGAAAACCATTTGGATA 735
Db 183 AsnProAlaArgSerPheAlaValLeuThrArgAsnPheThrAsnHisTrpVal 202

QY 736 TATTGGTGGGCCCATCATAGGAGCTGTCTCGTGTGGCTTTATGAGTATGTCTTCTTC 795
Db 203 TyrTrpValGlyProIleIleGlyAlaLeuGlyLeuValTyAspPheIleLeu 222

QY 796 TGTCCAGATGTT-----GAATTCAAACCTCGTTTTTAAAGAGCCTTTCAGCAA 843
Db 223 PheProArgMetArgGlyLeuSerGlu-ArgLeuSerIleLeuLysGlyAlaArgProAl 242

QY 844 GCTGCCAGCAACAAAGAGGAGCTTACA 871
Db 242 aGluProGluGlyGlnGlnAlaThr 251

```

RESULT 7

S53423  
major intrinsic protein (MIP26) - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 01-Aug-1995 #sequence\_revision 01-Sep-1995 #text\_change 09-Jul-2004  
C:Accession: S53423; S10491  
R:Dilsiz, N.; Crabbe, M.J.C.  
Biochem. J. 305, 753-759, 1995  
A>Title: Heterologous expression in Escherichia coli of native and mutant forms of the m  
A:Reference number: S53423; MUID:95151007; PMID:7848273  
A:Accession: S53423  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-261 <DIL>  
A:Cross-references: UNIPROT:P09011  
N:Kent, N.A.; Shiels, A.  
Nucleic Acids Res. 18, 4256, 1990  
A>Title: Nucleotide and derived amino-acid sequence of the major intrinsic protein of ra  
A:Reference number: S10491; MUID:90332430; PMID:2377471  
A:Accession: S10491  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-261 <KEN>  
A:Cross-references: EMBL:X53052; NID:g57621; PIDN:CAA37219.1; PID:g1334235  
C:Superfamily: lens fiber membrane major intrinsic protein

Alignment Scores:  
Pred. No.: 5,48e-42 Length: 261  
Score: 600.00 Matches: 108  
Percent Similarity: 68.47% Conservative: 44  
Best Local Similarity: 48.65% Mismatches: 66  
Query Match: 27.51% Indels: 4  
DB: 2 Gaps: 1

US-10-723-180-1 (1-1152) x S53423 (1-261)

QY 136 GCTTCTCGAAGACGTCACAGCGGAATTTCTGGCCATGCTTATTTTCTCTCTCAGC 195  
Db 6 SerPheTrpAlaIlePheAlaGluPheAlaThrLeuPheTyrValPheGly 25  
QY 196 CTGGATCCACCATCACTGGGTGGACAGAAAGCTTTACCGGTGACATGGTTCTC 255  
Db 26 LeuGlySerSerLeuArgTrpAla-----ProGlyProLeuHisValLeuGln 41  
QY 256 ATCTCCCTTGTGTTGGATCGATCGATTCGAACATGGTGCAGTCTTGGCCATTCAGC 315  
Db 42 ValAlaLeuAlaPheGlyLeuAlaLeuAlaThrLeuValGlnThrValGlyHisIleSer 61  
QY 316 GTGGCCACATCAACCTCGAGTGTGGCCATGGTGCACAGGAGATCAGCATC 375  
Db 62 GlyAlaHisValAsnProAlaValThrPheAlaPheLeuValGlySerGlnMetSerLeu 81  
QY 376 GCAAGTCTGTTCTTACATCGAGCCCGAGTGCCTGGGGCCATCATTTGAGCAGGAATC 435  
Db 82 LeuArgAlaPheCysTyrIleAlaAlaGlnLeuLeuGlyAlaValAlaGlyAlaVal 101  
QY 436 CTCATCTGGTTCACCTCCAGTGTGGTGGAGGCTGGAGTCCACATGTTTCATGA 495  
Db 102 LeuTyrSerValThrProProAlaValArgGlyAsnLeuAlaLeuAsnThrLeuHisAla 121  
QY 496 AATCTTACCGCTGTCATGGTCTCCTGTTGATGATTAATCACAATTTCAATTTGGTGT 555  
Db 122 GlyValSerValGlyGlnAlaThrThrValGluIlePheLeuThrLeuGlnPheValLeu 141  
QY 556 ACTATCTTTGCCAGCTGTGATTCACAAACGAGCTGATGTCACTGGCTCAATAGCTTTAGCA 615  
Db 142 CysIlePheAlaThrTyrAspGluArgAsnGlyArgMetGlySerValAlaLeuAla 161  
QY 616 ATGGATTTTCTGTCGATTCGACATTTATTTGCAATCAATATATCTGTTCCAGCATG 675  
Db 162 ValGlyPheSerLeuThrLeuGlyHisLeuPheGlyMetTyrThrGlyAlaGlyMet 181  
QY 676 AATCCGCGCCGATCTCTTGGACCTGCAGTTATCATGGGAAATTTGGAAACCATTTGGATA 735  
Db 44 IleAlaMetAlaPheGlyLeuGlyIleGlyThrLeuValGlnAlaLeuGlyHisIleSer 63

Db 182 AsnProAlaArgSerPheAlaProAlaIleLeuThrArgAsnPheSerAsnHisTrpVal 201  
QY 736 TATTGGTGGCCCATCATAGAGCTGTCTCGCTGGTGGCTTTATGAGTATGTCTTC 795  
Db 202 TyrTrpValGlyProIleIleGlyGlyLeuGlySerLeuLeuTyrAspPheLeuLeu 221  
QY 796 TGTCCA 801  
Db 222 PhePro 223

RESULT 8  
A53442  
aquaporin 2 - human  
N:Alternate names: water-channel protein  
C:Species: Homo sapiens (man)  
C>Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 09-Jul-2004  
C:Accession: A53442; I55410; I51877  
R:Dean, P.M.T.; Verdijk, M.A.J.; Knoers, N.V.A.M.; Wieringa, B.; Monnens, L.A.H.; van Os  
Science 264, 92-95, 1994  
A>Title: Requirement of human renal water channel aquaporin-2 for vasopressin-dependent  
A:Reference number: A53442; MUID:94188717; PMID:8140421  
A:Accession: A53442  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-271 <DEB>  
A:Cross-references: UNIPROT:P41181; GB:Z29491; NID:g474058; PIDN:CAA82627.1; PID:g474059  
R:Uchida, S.; Sasaki, S.; Fushimi, K.; Marumo, F.  
J. Biol. Chem. 269, 23451-23455, 1994  
A>Title: Isolation of human aquaporin-CD gene.  
A:Reference number: I55410; MUID:94375443; PMID:7522228  
A:Accession: I55410  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-271 <UCH>  
A:Cross-references: GB:D31846; NID:g567249; PIDN:BAA06632.1; PID:g567250  
R:van Lieburg, A.F.; Verdijk, M.A.; Knoers, V.V.; van Essen, A.J.; Proemans, W.; Mallma  
Am. J. Hum. Genet. 55, 648-652, 1994  
A>Title: Patients with autosomal nephrogenic diabetes insipidus homozygous for mutations  
A:Reference number: I51877; MUID:95029363; PMID:7524315  
A:Accession: I51877  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-186, 'C', 188-271 <LIE>  
A:Cross-references: GB:S73196; NID:g685000; PIDN:AAB31999.1; PID:g685001  
C:Genetics:  
A:Gene: GDB:AQP2  
A:Cross-references: GDB:141853; OMIM:107777  
A:Map position: 12q13-12q13  
A:Intron: 120/3; 175/3; 202/3  
C:Superfamily: lens fiber membrane major intrinsic protein  
C:Keywords: transmembrane protein; water channel

Alignment Scores:  
Pred. No.: 9,79e-42 Length: 271  
Score: 597.00 Matches: 114  
Percent Similarity: 69.37% Conservative: 40  
Best Local Similarity: 51.35% Mismatches: 64  
Query Match: 27.37% Indels: 4  
DB: 2 Gaps: 2

US-10-723-180-1 (1-1152) x A53442 (1-271)

QY 136 GCTTCTGGAAGACGTCACAGCGGAATTTCTGGCCATGCTTATTTTGTCTCTCAGC 195  
Db 8 AlaPheSerArgAlaValPheAlaGluPheLeuAlaThrLeuLeuPheValPheGly 27  
QY 196 CTGGATCCACCATCACTGGGTGGACAGAAAGCTTTACCGGTGACATGGTTCTC 255  
Db 28 LeuGlySerAlaLeuAsnTrp-----ProGlnAlaLeuPro---SerValLeuGln 43  
QY 256 ATCTCCCTTGTGTTGACCTGACATTCACACCATGGTGCAGTGTGTTGGCCATTCAGC 315  
Db 44 IleAlaMetAlaPheGlyLeuGlyIleGlyThrLeuValGlnAlaLeuGlyHisIleSer 63



Best Local Similarity: 43.57% Mismatches: 77  
Query Match: 27.03% Indels: 31  
DB: 2 Gaps: 5

US-10-723-180-1 (1-1152) x I51164 (1-272)

QY 133 CAAGCTTTCTGGAAGAGTCACAGCGGAATTTCTGGCCATCTTATTTTGTCTCTC 192  
DB 8 LysAlaPheTrpArgAlaValIleAlaGluPheAlaMetIleuPheValPheIle 27

QY 193 AGCTGGATCCACCATCACTGGGTGGACAGAAAAGCCCTTTACCGTC----- 243  
DB 28 SerIleGlyAlaAlaLeuGlyPheAen-----PheProIleGluGlyLys 42

QY 244 -----GACATGGTTCTCATCTCCCTTTGCTTTGGATC 276  
DB 43 AlaAenGlnThrValGlyArgSerGlnAspIleValSerLeuAlaPheGlyIle 62

QY 277 AGCATTTGCAACCATCGTCAGTCTTGGCCATATCAGCGTGGCCACATCAACCTGCA 336  
DB 63 SerIleAlaThrMetAlaGlnSerValGlyHisValSerGlyAlaHisLeuAenProAla 82

QY 337 GTGACTGTGGCCATCGTGTGCACAGGAAGATCAGCATCGCCAAAGTCTGTCTTACATC 396  
DB 83 ValThrLeuGlyCysLeuLeuSerCysGlnIleSerIleLeuLysAlaValMetTyrIle 102

QY 397 GCAGCCCATGCTGGGCCATCATTTGGAGCAGGAATCTCTATCTGTGTACACCTCC 456  
DB 103 IleAlaGlnCysLeuGlyAlaValAlaThrAlaIleLeuSerGlyIleThrSerGly 122

QY 457 AGTGTGTGGGAGGCTGGGAGTCACCATGGTTCATGGAAATCTTACCGCTGTCTATGGT 516  
DB 123 LeuGluAenAenSerLeuGlyLeuAenglyLeuSerProGlyValSerAlaGlyGlnGly 142

QY 517 CTCCTGGTGTGATGATCAATCAATTTCAATTTGGTGTCTTACTATCTTTGCCAGCTGTGAT 576  
DB 143 LeuGlyValGluIleLeuValThrPheGlnLeuValLeuCysValAlaValThrAsp 162

QY 577 TCCAAACGGATGATCTACTGGCTCAATAGCTTTAGCAATTTGGATTTCTGTGCAATT 636  
DB 163 ArgArgArgHisAspValSerGlySerValProLeuAlaIleGlyLeuSerValAlaLeu 182

QY 637 GGACATTTATTTGCAATCAATATATCTGTGTGCAGCATGAATCCCGCCGATCTTTGGA 696  
DB 183 GlyHisLeuIleAlaIleAspPyrThrGlyCysGlyMetAenProAlaArgSerPheGly 202

QY 697 CTTGAGTTATCATGGAAATTTGGAAACCATTCGATATATTTGGTGGGCCCATCAT 756  
DB 203 SerAlaValLeuThrLysAenPheThrTyrHisTrpIlePheThrValGlyProMetIle 222

QY 757 GGAGCTGTCTCGTGGTGGCTTTATGAGTATGCTCTGTCTCCAGATGTTGAATTCAAA 816  
DB 223 GlyGlyAlaAlaAlaIleIleTyrAspPheIleLeuAlaPro----- 237

QY 817 CGTCTGTTTAAAGAGCTTTCAGCAAGCTGCCAGCAACAAAGGAAGCTACATGGAG 876  
DB 238 ---ArgThrSerAspLeuThrAspArgMetLysValTrpThrAsnGlyGlnValGlu 256

QY 877 GTGGAG-----GACAACAGGAGTCAGGTAGAGCGGATGACCTGATTTCTAAACCT 927  
DB 257 TyrGluLeuAspGlyAspAenThrArgValGlu-----MetLysPro 271

RESULT 11  
164818  
water-channel aquaporin 2 - human  
C:Species: Homo sapiens (man)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 13-Aug-1999  
C:Accession: I64818  
R:van Lieburg, A.F.; Verdijk, M.A.; Knoers, V.V.; van Esen, A.J.; Proemans, W.; Mallma  
Am. J. Hum. Genet. 55, 648-652, 1994  
A:Title: Patients with autosomal nephrogenic diabetes insipidus homozygous for mutations  
A:Reference number: I51877; MUID:95029363; PMID:7524315  
A:Accession: I64818

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-271 <RES>  
A:Cross-references: GB:S73197; NID:G684998; PIDN:AAB31998.1; PID:G684999  
C:Genetics:  
A:Gene: GDB:AQP2  
A:Cross-references: GDB:141853; OMIM:107777  
A:Map position: 12q13-12q13  
C:Superfamily: lens fiber membrane major intrinsic protein

Alignment Scores:  
Pred. No.: 4,56e-41 Length: 271  
Score: 589.00 Matches: 113  
Percent Similarity: 68.92% Conservative: 40  
Best Local Similarity: 50.90% Mismatches: 65  
Query Match: 27.01% Indels: 4  
DB: 2 Gaps: 2

US-10-723-180-1 (1-1152) x 164818 (1-271)

QY 136 GCTTTCTGGAAGAGTCACAGCGGAATTTCTGGCCATGCTTATTTTGTCTCTCAGC 195  
DB 8 AlaPheSerArgAlaValPheAlaGluPheLeuAlaThrLeuLeuPheValPhePheGly 27

QY 196 CTGGGATCCACCATCACTGGGTGGACAGAAAAGCCCTTTACCGTTCGACATGGTCTC 255  
DB 28 LeuGlySerAlaLeuAenTrp-----ProGlnAlaLeuPro---SerValLeuGln 43

QY 256 ATCTCCCTTTCTTGGACTCAGCATGCAACCATGTCAGTGTGGCCATATCAGC 315  
DB 44 IleAlaMetAlaPheGlyLeuGlyIleGlyThrLeuValGlnAlaLeuGlyHisIleSer 63

QY 316 GGTGCCCATCAACCTGTGAGTGTGGCCATGTCACAGGAAGATCAGCATC 375  
DB 64 ArgAlaHisIleAenProAlaValThrValAlaCysLeuValGlyCysHisValSerVal 83

QY 376 GCCAAGTCTGTCTTACATCGCAGCCAGTGTGGGGCCATCATTTGGAGCAGGAATC 435  
DB 84 LeuArgAlaAlaPheTyrValAlaGlnLeuLeuGlyAlaValAlaGlyAlaAlaLeu 103

QY 436 CTCTATCTGTGTCCACCTCCCGATGTGTGGAGGCTGGGAGTCACCATGGTTCATGGA 495  
DB 104 LeuHisGluIleThrProAlaAspIleArgGlyAspLeuAlaValAenAlaLeuSerAen 123

QY 496 AATCTTACCGTGTGTGTCATGTCCTGTGTGAGTTGATATACATTTCAATTTGGTGTTC 555  
DB 124 SerThrThrAlaGlyGlnAlaValThrValGluLeuPheLeuThrLeuGlnLeuValLeu 143

QY 556 ACTATCTTGGCCAGCTGTGATTCCAAAGCGGACTGATGTCCTGGCTCAATAGCTTTAGCA 615  
DB 144 CysIlePheAlaSerThrAspGluArgGlyGluAenProGlyThrProAlaLeuSer 163

QY 616 ATTGGATTTCTGTGCAATTTGGACATTTATTTGCAATTAATTTACTGTGGCCAGCATG 675  
DB 164 IleGlyPheSerValAlaLeuGlyHisLeuLeuGlyIleHisTyrThrGlyCysSerMet 183

QY 676 AATCCGCCCATCTTTGGACCTGTGATTCATCATGGAAATTTGGAAAACATTGGATGATA 735  
DB 184 AenProAlaArgSerLeuAlaProAlaValThrGlyLysPheAspHisTrpVal 203

QY 736 TATTGGGTGGGCCCATCATAGGAGCTGCTCGCTGGTGGCCCTTATGATGATGTCTTC 795  
DB 204 PheTrpIleGlyProLeuValGlyAlaIleLeuGlySerLeuLeuTyrAsnTyrValLeu 223

QY 796 TGTCCA 801  
DB 224 PhePro 225

RESULT 12  
B4499  
major intrinsic protein homolog DER2 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004



C:Accession: B44499  
R:lanahan, A.; Williams, J.B.; Sanders, L.K.; Nathans, D.  
Mol. Cell. Biol. 12, 3919-3929, 1992  
A:Title: Growth factor-induced delayed early response genes.  
A:Reference number: A44499; MUID:92375060; PMID:1508193  
A:Accession: B44499  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-269 <LAN>  
A:Cross-references: UNIPROT:Q02013; GB:L02914; NID:g193642; PIDN:AAB53928.1; PID:g193643  
A:Experimental source: BALB/c 3T3 cells  
A>Note: sequence extracted from NCBI backbone (NCBIP:111645)  
C:Superfamily: lens fiber membrane major intrinsic protein  
C:Keywords: transmembrane protein

Alignment Scores:  
Pred. No.: 9,86e-40 Length: 269  
Score: 573.00 Matches: 123  
Percent Similarity: 63.43% Conservative: 47  
Best Local Similarity: 45.90% Mismatches: 84  
Query Match: 26.27% Indels: 14  
DB: 2 Gaps: 6

US-10-723-180-1 (1-1152) x B44499 (1-269)

QY 139 TTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTTTATTTTCTCTCTCAGCCTG 198  
Db 10 PheTrpArgAlaValAlaGluPheLeuAlaMetThrLeuPheValPheIleSerIle 29

QY 199 GGATCCACCATCACTGGGGT-----GGACAGAAAGCCCTTTACCGTCGACATG 249  
Db 30 GlySerAlaLeuGlyPheAsnTyrProLeuGluArgAsnGlnThrLeuValGlnAspAsn 49

QY 250 GTTCTCATCTCCCTTGTCTTGGACTCAGCATTCAGACCATGTCGTCGCTTTGGCCAT 309  
Db 50 ValIleValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis 69

QY 310 ATCAGCGGTGGCCACATCAACCTGTCAGTGACTGTGGCCATGTCGTCACAGGAGATC 369  
Db 70 IleSerGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeuLeuSerCysGlnIle 89

QY 370 AGCATGCCCAAGTCGTCTTCTACATCGCAGCCAGTCGCTGGGGCCATCAATCGGACA 429  
Db 90 SerIleLeuArgAlaValMetTyrIleIleAlaGlnCysValGlyAlaIleValAlaThr 109

QY 430 GGAATCCTCTATCTGTGTACACCTCCAGCTGTGGTGGAGGCTGGGATCACC---ATG 486  
Db 110 AlaIleLeuSerGlyIleThrSerSerLeuValAlaAspAsnSerLeuGlyArgAsnAspLeu 129

QY 487 GTTCATGGAATCTTACCCTGTCATGTCCTCTGTTGAGTTGATGATATCATCTTCAA 546  
Db 130 AlaHisGly--ValAsnSerGlyGlnGlyLeuGlyIleGluIleIleGlyThrLeuGln 148

QY 547 TTGGTGTTCATCTCTTCCAGCTGTGTATCCAAACGACATGATGTCTCGCTCAATA 606  
Db 149 LeuValLeuCysValLeuAlaThrThrAspArgArgArgAspLeuGlyGlySerAla 168

QY 607 GCTTTAGCAATGGATTTCTCTGCAATTGACATTTATTTGCAATCATTAATCTGGT 666  
Db 169 ProLeuAlaIleGlyLeuSerValAlaLeuGlyHisLeuLeuAlaIleAspTyrThrGly 188

QY 667 GCCAGCATGAATCCCGCCGATCTTTGGACCTCGAGTTATCATGGGAATTTGGGAAAC 726  
Db 189 CysGlyIleAsnProAlaArgSerPheGlySerAlaValLeuThrArgAsnPheSerAsn 208

QY 727 CATTTGATATATGGTTGGGCCCATCATPAGAGCTGTCTCGCTGGTGGCCCTTTATGAG 786  
Db 209 HisTrpIlePheTrpValGlyProPheIleGlyGlyAlaLeuAlaValLeuIleTyrAsp 228

QY 787 TATGTCTTCTGTCCA---CATGTTGAAATTCACACGTCGCTTTTAAAGAGCCCTTCAGCAA 843  
Db 229 PheIleLeuAlaProArgSerSerAspPheThrAspArgMetLeuValTrpThrSerGly 248

QY 844 GCTGCCCGACCAACAAAGGAAGCTACATGGAGGTGGAGGACACACAGGAGTCAGGTAG 903  
Db 249 GlnValGluGluTyrAsp-----LeuAspAlaAspAspIleAsnSerArgValGlu 265

QY 904 ACGGATGACCTGATCTCTAAACCT 927  
Db 266 -----MetLysPro 268

RESULT 13  
JC1320  
water channel protein CHIP28 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text\_change 09-Jul-2004  
A:Accession: JC1320; JT0749; S37639; A44395  
R:Deen, P.M.T.; Dempster, J.A.; Wieringa, B.; Van Os, C.H.  
Biochem. Biophys. Res. Commun. 188, 1267-1273, 1992  
A:Title: Isolation of a cDNA for rat CHIP28 water channel: High mRNA expression in kidney  
A:Reference number: JC1320; MUID:93075224; PMID:1280133  
A:Accession: JC1320  
A:Molecule type: mRNA  
A:Residues: 1-269 <DEB>  
A:Cross-references: UNIPROT:P29975; GB:X67948; GB:S49827; NID:g55941; PIDN:CAA48134.1; F  
A:Experimental source: kidney  
R:Ma, T.; Frigeri, A.; Skach, W.; Verkman, A.S.  
Biochem. Biophys. Res. Commun. 197, 654-659, 1993  
A:Title: Cloning of a novel rat kidney cDNA homologous to CHIP28 and WCH-CD water channel  
A:Reference number: JT0749; MUID:94092144; PMID:7505572  
A:Accession: JT0749  
A:Molecule type: mRNA  
A:Residues: 1-4,'P',6-121,'SN',124-269 <MAT>  
A:Experimental source: kidney  
R:Shanahan, C.M.; Weisberg, P.L.; Metcalfe, J.C.  
Circ. Res. 73, 193-204, 1993  
A:Title: Isolation of gene markers of differentiated and proliferating vascular smooth m  
A:Reference number: S37637; MUID:93284726; PMID:8508530  
A:Accession: S37639  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2,'T',4-158,'O',160-269 <SHA>  
A:Cross-references: EMBL:X71069; NID:g313803; PIDN:CAA50395.1; PID:g313804  
R:Zhang, R.; Skach, W.; Hasegawa, H.; van Hoek, A.N.; Verkman, A.S.  
J. Cell Biol. 120, 359-369, 1993  
A:Title: Cloning, functional analysis and cell localization of a kidney proximal tubule  
A:Reference number: A44395; MUID:93132068; PMID:8421053  
A:Accession: A44395  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-4,'P',6-269 <ZHA>  
A:Cross-references: GB:X70257; GB:S52660; NID:g312923; PIDN:CAA49761.1; PID:g312924  
A:Experimental source: renal cortex  
A>Note: sequence extracted from NCBI backbone (NCBIP:122674)  
C:Superfamily: lens fiber membrane major intrinsic protein  
C:Keywords: glycoprotein; transmembrane protein  
F:42,205/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:  
Pred. No.: 1,2e-39 Length: 269  
Score: 572.00 Matches: 120  
Percent Similarity: 63.67% Conservative: 50  
Best Local Similarity: 44.94% Mismatches: 85  
Query Match: 26.23% Indels: 12  
DB: 2 Gaps: 4

US-10-723-180-1 (1-1152) x JC1320 (1-269)

QY 139 TTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTTTATTTTCTCTCAGCCTG 198  
Db 10 PheTrpArgAlaValAlaGluPheLeuAlaMetThrLeuPheValPheIleSerIle 29

QY 199 GGATCCACCATCACTGGGGT-----GGACAGAAAGCCCTTTACCGTCGACATG 249  
Db 30 GlySerAlaLeuGlyPheAsnTyrProLeuGluArgAsnGlnThrLeuValGlnAspAsn 49



```

QY 250 GTTCTCATCTCCCTTGTCTTGGACTCAGCATTTGACCAACCATCGTGCAGTGTCTTGGCCAT 309
Db 50 VallysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis 69
QY 310 ATCAGCGGTGGCCACATCAACCTCGCAGTACTGTGGCCATGTGTGCACACGAGGATC 369
Db 70 IleserGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeuSerCysGlnIle 89
QY 370 AGCATGCCAAGCTGTCTTACATCGCAGCCCGAGTGCCTGGGGCCCATCAATGGAGCA 429
Db 90 SerIleLeuArgAlaValMetTyrIleIleAlaGlnCysValGlyAlaIleValAlaSer 109
QY 430 GGAATCCCTATCTGGTCACACTCCAGTGTGGTGGAGGCTGGGAGTCCACATGTT 489
Db 110 AlaIleLeuSerGlyIleThrSerSerLeuLeuGluAenSerLeuGlyArgAsnAspLeu 129
QY 490 CATGAAATCTTACCCTGGTGCATGCTCTCTGTTGAGTTGATTAATACATTTCAATTG 549
Db 130 AlaArgGlyValAsnSerGlyGlnGlyLeuGlyIleGluIleIleGlyThrLeuGlnLeu 149
QY 550 GTGTTTATCTATCTTCCAGCTGTGATTCACAAACGAGTGTGATGTCACCTGGCTCAATAGCT 609
Db 150 ValLeuCysValLeuAlaThrThrAspArgArgAspLeuGlyGlySerAlaPro 169
QY 610 TTAGCAATTTGATTTCTGTGCAATTGGACATTTATTGCAATTAATTAATGCTGCC 669
Db 170 LeuAlaIleGlyLeuSerValAlaLeuGlyHisLeuLeuAlaIleAspTyrThrGlyCys 189
QY 670 AGCATGAATCCCGCCGATCTCTTGGACCTGCAGTATCATCGGAAATTTGGAAACCAT 729
Db 190 GlyIleAsnProAlaArgSerPheGlySerAlaValLeuThrArgAsnPheSerAsnHis 209
QY 730 TGGATATATTTGGTGGGCCCATCATAGAGCTGTCTCGCTGGTGGCTTTATGATGAT 789
Db 210 TrpIlePheTrpValGlyProPheIleGlySerAlaLeuAlaValLeuIleTyrAspPhe 229
QY 790 GTCCTCTGTCCA--GATGTTGAATTCAACTCGCTTTTAAGAGAGCTTCAGCAAGCT 846
Db 230 IleLeuAlaProArgSerSerAspPheThrAspArgMetLysValTrpThrSerGlyGln 249
QY 847 GCCCAGCAACAAAGAGAGCTTACATGGAGGTGGAGGACCAACAGCAGTACGAGTACG 906
Db 250 ValGluGluTyrAsp-----LeuAspAlaAspAspIleAsnSerArgValGlu--- 265
QY 907 GATGACCTGATCTTAAACCT 927
Db 266 -----MetLysPro 268

RESULT 14
JC2348
water channel protein CHIP29 - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: JC2348
R:Patil, R.V.; Yang, X.; Saico, I.; Coca-Prados, M.; Wax, M.B.
Biochem. Biophys. Res. Commun. 204, 861-866, 1994
A:Title: Cloning of a novel cDNA homologous to CHIP28 water channel from ocular ciliary
A:Reference number: JC2348; MUID:95071406; PMID:7526855
A:Accession: JC2348
A:Keywords: glycoprotein
A:Molecule type: mRNA
A:Residues: 1-271 <PAT>
A:Cross-references: UNIPROT:P47865; GB:S74759; NID:g786448; PIDN:AAB32365.1; PID:g786450
A:Experimental source: ciliary epithelium
A:Superfamily: lens fiber membrane major intrinsic protein
C:Keywords: glycoprotein
F:4/2/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2,138-39 Length: 271
Score: 569.00 Matches: 118
Percent Similarity: 61.84% Conservative: 52
Best Local Similarity: 42.91% Mismatches: 79
Query Match: 26.09% Indels: 26

```

```

DB: 2 Gaps: 5
US-10-723-180-1 (1-1152) x JC2348 (1-271)
QY 139 TTCTGGAAGCAGTCACAGCGGAATTTCTGGCCATGCTTATTTTGTCTCTCTCAGCCTG 198
Db 10 PheTrpArgAlaValAlaGluPheLeuAlaMetIlePheLeuPheIleSerIle 29
QY 199 GGATCCACCATCAATGGGTGGAAACAGAAAGCCTTTACCG----- 240
Db 30 GlySerAlaLeu-----GlyPheHisTyrProIleLysSerAsnGlnThrThrGly 46
QY 241 -----GTCGACATGGTTCTCATCTCCCTTTGCTTTGGACTCAGCATTTGCAACCATG 294
Db 47 AlaValGlnAspAsnValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAla 66
QY 295 CAGTGTCTTGGCCATATCAGGGGTGGCCACATCAACCTCGCAGTACGTGCGCCATGCTG 354
Db 67 GlnSerValGlyHisIleSerGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeu 86
QY 355 TGCAACAGGAAGATCAGCATCGCCCAAGTCTCTTACATCGCAGCCAGTGCCTGGGG 414
Db 87 LeuSerCysGlnIleSerValLeuArgAlaIleMetTyrIleIleAlaGlnCysValGly 106
QY 415 GCCATCATTTGAGCAGGAAATCTCTATCTGTGTGCACCTCCAGTGTGTGGAGGCTG 474
Db 107 AlaIleValAlaThrAlaIleLeuSerGlyIleThrSerSerLeuProAspAsnSerLeu 126
QY 475 GGAGTCACCATGGTTTCATGGAAATTTACCTCGTGCATGCTCTCTCGTTGAGTTGATA 534
Db 127 GlyLeuAsnAlaLeuAlaProGlyValAsnSerGlyGlnGlyLeuGlyIleGluIle 146
QY 535 ATCATATTCAATTTGGTGTCTTACTATCTTCCAGCTGTGTATCCAAACGAGCATGATGTC 594
Db 147 GlyThrLeuGlnLeuValLeuCysValLeuAlaThrThrAspArgArgAspLeu 166
QY 595 ACTGGCTCAATAGCTTTAGCAATTTGGATTTCTGTGCAATTGGACATTTATTGCAATC 654
Db 167 GlyGlySerGlyProLeuAlaIleGlyPheSerValAlaLeuGlyHisLeuAlaIle 186
QY 655 AATTATATCTGTGTCAGCATCAATCCCGCCGATCTCTTGGACCTGCAGTTCATCGGA 714
Db 187 AspTyrThrGlyCysGlyIleAsnProAlaArgSerPheGlySerSerValIleThrHis 206
QY 715 AATTGGGAAACCATTTGGATATATTTGGTGTGGGCCCATCATAGAGCTGTCTCGCTGCT 774
Db 207 AsnPheGlnAspHisTrpIlePheTrpValGlyProPheIleGlyAlaAlaLeuAlaVal 226
QY 775 GGCCTTTATGATGATGCTCTCTGTCCAGATGTTGAATTCAAACGCTGTTTAAAGAGGCC 834
Db 227 LeuIleTyrAspPheIleLeuAlaPro-----ArgSerSerAspLeu 240
QY 835 TTCAGCAAGCTGCCAGCAACAAAGAAAGAGTAC-----ATCGAGGTGGAG 882
Db 241 ThrAspArgValLysValTrpThrSerGlyGlnValGluGluTyrAspLeuAspAlaAsp 260
QY 883 GACAAACAGGACTCAGGTAGAGACCGATGACCTGATTTCTTAAACCT 927
Db 261 AspIleAsnSerArgValGlu-----MetLysPro 270

RESULT 15
152366
uterine water channel - human
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
C:Accession: I52366
R:Li, X.; Yu, H.; Koide, S.S.
Biochem. Mol. Biol. Int. 32, 371-377, 1994
A:Title: The water channel gene in human uterus.
A:Reference number: I52366; MUID:94290349; PMID:7517253
A:Accession: I52366
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

```

A;Residues: 1-269 <RES>  
A;Cross-references: GB:S73482; NID:G688357; PIDN:AAB31193.1; PID:G688358  
C;Superfamily: lens fiber membrane major intrinsic protein

Alignment Scores:  
Pred. No.: 6,74e-39 Length: 269  
Score: 563.00 Matches: 118  
Percent Similarity: 61.11% Conservative: 47  
Best Local Similarity: 43.70% Mismatches: 87  
Query Match: 25.81% Indels: 18  
DB: 2 Gaps: 4

US-10-723-180-1 (1-1152) x IS2366 (1-269)

QY 139 TTCTGGAAGCAGTCACAGCGGAATTTCTGGCCATGCTTATTTTTGTCTCTCCTCAGCCTG 198  
Db 10 PheTrpArgAlaValValAlaGluPheLeuAlaThrThrLeuPheValPheIleSerIle 29  
QY 199 GGNATCCACCATCAACTGGGGT-----GGACAGAAAGCCTTTACCGGTGCGACATG 249  
Db 30 GlySerAlaLeuGlyPheLysTyrProValGlyAsnAsnGlnThrValValGlnAspAsn 49  
QY 250 GTTCTCATCTCCCTTTGCTTTGGACTCAGCATTCACCATGCTGTCAGTGTCTTTGGCCAT 309  
Db 50 ValLysValSerLeuAlaPheGlyLeuSerIleAlaThrLeuAlaGlnSerValGlyHis 69  
QY 310 ATCAGCGGTGGCCACATCAACCTGCAGTGACTGTGGCCATGGTGTGCACACAGGAAGATC 369  
Db 70 IleSerGlyAlaHisLeuAsnProAlaValThrLeuGlyLeuLeuLeuSerCysGlnIle 89  
QY 370 AGCATGCCAAGTCTCTTCTACATCGCAGGCCAGTGTCTGGGGCCATCATTTGGAGCA 429  
Db 90 SerIlePheArgAlaLeuMetTyrIleIleAlaGlnCysValGlyAlaIleValAlaThr 109  
QY 430 GGNATCCTCTATCTGTGTACACTCCACGTGTGTGGAGGCTGGAGTCCACCTGGTT 489  
Db 110 AlaIleLeuSerGlyIleThrSerLeuThrGlyAsnSerLeuGlyArgAsnAspLeu 129  
QY 490 CATGGAATCTTACCCTGCTCATGCTCTCTGTTGAGTGTGATATCATCATTTCAATTG 549  
Db 130 AlaAspGlyValAsnSerGlyGlnGlyLeuGlyIleGluIleIleGlyThrLeuGlnLeu 149  
QY 550 GTGTTTACTATCTTCCAGCTGTGTATCCAAACGGACTGATGTCTACTGGCTCAATAGCT 609  
Db 150 ValLeuCysValLeuAlaThrThrAspArgArgAspLeuGlyGlySerAlaPro 169  
QY 610 TTAGCAATTGGATTTCTCTGCAATTGGACATTTATTTGCAATCAATTATCTGTGTC 669  
Db 170 LeuAlaIleGlyLeuSerValAlaLeuGlyHisLeuLeuAlaIleAspTyrThrGlyCys 189  
QY 670 AGCATGAATCCGCGCCGATCCTTTGGACCTGCAGTTATCATGGGAAATTTGGGAAACCAT 729  
Db 190 GlyIleAsnProAlaArgSerPheGlySerAlaValIleThrHisAsnPheSerAsnHis 209  
QY 730 TGGATATATTGGTGGGCCCATCATAGAGAGCTGCTCTGCTGGTGGCTTTTATGAGTAT 789  
Db 210 TrpIlePheTrpValGlyProPheIleGlyAlaLeuAlaValLeuIleTyrAspPhe 229  
QY 790 GTCTTCTGTCCAGATTTGAATTCAACGCTGTTTAAAGAGCCTTCAGCAAGCTGCC 849  
Db 230 IleLeuAlaPro-----ArgSerAspLeuThrAspArgValLys 243  
QY 850 CAGCAAAACAAAGGAAGCTAC-----ATGGAGGTGGAGGACAAACAGAGTCTAG 897  
Db 244 ValTrpThrSerGlyGlnValGluGluTyrAspLeuAspAlaAspAspIleAsnSerArg 263  
QY 898 GTAGAGCGGATGACCTGATCTTCTAAACCT 927  
Db 264 ValGlu-----MetLysPro 268

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 8, 2005, 11:55:12 ; Search time 229 Seconds  
(without alignments)  
5152.098 Million cell updates/sec

Title: US-10-723-180-1  
Perfect score: 2181  
Sequence: 1 999gcaggcaatgagagctg.....attactcgtctaaacaata 1152

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-O=/cgn2.1/USPTO spool\_p/US10723180/runat\_08092005\_123256\_29239/app\_query.fasta\_1.1351  
-DB=UniProt -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10723180 @CNC 1.1.248 @runat\_08092005\_123256\_29239 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_treml.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1695	77.3	323	1 AQP4_HUMAN	P55087 homo sapien
2	1628	74.6	323	1 AQP4_BOVIN	O77750 bos taurus
3	1623	74.4	323	2 Q866S4	Q86684 ovis aries
4	1596	73.2	324	1 AQP4_DIPME	Q923j4 dipodomys m
5	1595	73.1	323	1 AQP4_RAT	P47863 rattus norv
6	1567	71.8	323	1 AQP4_MOUSE	P55088 mus musculus
7	1505	69.0	301	2 Q6XVT6	Q6XVT6 ovis aries
8	1466	67.2	335	2 Q8UVB8	Q8UVB8 coturnix co
9	1449	66.4	335	2 Q65YQ3	Q65YQ3 gallus gall
10	1103	50.6	224	2 Q8BR89	Q8BR89 mus musculus
11	1080.5	49.5	218	2 Q6L7A0	Q6L7A0 homo sapien
12	1037.5	47.6	320	2 Q6AZD2	Q6AZD2 brachydanio
13	884	40.5	178	2 Q46426	Q46426 oryctolagus
14	707	32.4	149	2 Q8K4M1	Q8K4M1 mus musculus
15	656.5	30.1	249	1 AQP_AEDAE	Q9nhw7 aedes aegypt
16	625	28.7	244	2 Q7PWV1	Q7PWV1 anopheles g

17	621	28.5	274	2 Q6T6Z9	Q6T6Z9 coturnix co
18	620	28.4	271	2 Q9YI68	Q9YI68 bufo marinu
19	617	28.3	265	1 AQP5_HUMAN	P55064 homo sapien
20	615	28.2	299	2 Q6AYU6	Q6AYU6 rattus norv
21	614	28.2	265	1 AQP5_MOUSE	Q9WY4 mus musculus
22	614	28.2	265	1 AQP5_RAT	P47864 rattus norv
23	614	28.2	271	2 Q8UJR3	Q8UJR3 hyala japoni
24	612	28.1	263	1 MIP_HUMAN	P30301 homo sapien
25	606	27.8	263	1 MIP_BOVIN	P06624 bos taurus
26	605	27.7	271	1 AQP2_SHEEP	P062735 ovis aries
27	605	27.7	273	2 Q8AWH9	Q8AWH9 xenopus lae
28	604	27.7	263	2 Q6J8I9	Q6J8I9 ovis aries
29	604	27.7	271	1 AQP2_MOUSE	P56402 mus musculus
30	603.5	27.7	270	2 Q66LN3	Q66LN3 coturnix co
31	603.5	27.7	274	2 Q6IP27	Q6IP27 xenopus lae
32	602.5	27.6	268	2 Q9YI69	Q9YI69 bufo marinu
33	602	27.6	263	1 MIP_RANPI	Q06019 rana pipien
34	600	27.5	261	1 MIP_RAT	P09011 rattus norv
35	599	27.5	262	2 Q8QFW8	Q8QFW8 gallus gall
36	598.5	27.4	274	2 Q6DJ01	Q6DJ01 xenopus tro
37	597.5	27.4	263	1 MIP_MOUSE	P51180 mus musculus
38	597	27.4	271	1 AQP2_HUMAN	P41181 homo sapien
39	596.5	27.3	263	2 Q640N8	Q640N8 mus musculus
40	596.5	27.3	268	2 Q76E78	Q76E78 hyala japoni
41	595	27.3	271	2 Q8UJR4	Q8UJR4 hyala japoni
42	595	27.3	274	2 Q9YI67	Q9YI67 bufo marinu
43	594	27.2	271	1 AQP2_RAT	P34080 rattus norv
44	594	27.2	271	2 Q6FGT3	Q6FGT3 homo sapien
45	589.5	27.0	272	1 AQP4_RANES	P50501 rana esculle

#### ALIGNMENTS

RESULT 1  
AQP4\_HUMAN  
ID AQP4\_HUMAN STANDARD; PRT; 323 AA.  
AC P55087; P78564;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Aquaporin 4 (WCHA) (Mercurial-insensitive water channel) (MIWC).  
GN Name=AQP4;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal brain;  
RX MEDLINE=96032721; PubMed=7559426; DOI=10.1074/jbc.270.39.22907;  
RA Yang B., Ma T., Verkman A.S.;  
RT "cDNA cloning, gene organization, and chromosomal localization of a human mercurial insensitive water channel. Evidence for distinct transcriptional units.";  
RL J. Biol. Chem. 270:22907-22913 (1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9601457; PubMed=8601457; DOI=10.1016/0014-5793(96)00092-0;  
RA Misaka T., Abe K., Iwabuchi K., Kusakabe Y., Ichinose M., Miki K., Emori Y., Arai S.;  
RT "A water channel closely related to rat brain aquaporin 4 is expressed in acid- and pepsinogen-secretory cells of human stomach.";  
RL FEBS Lett. 381:208-212 (1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97008105; PubMed=8855281; DOI=10.1073/pnas.93.20.10908;  
RA Lu M., Lee M.D., Smith B.L., Jung J.S., Agre P., Verdijk M.A.J., Merckx G., Rijes J.P.L., Deen P.M.T.;  
RT "The human AQP4 gene: definition of the locus encoding two water channel polypeptides in brain.";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:10908-10912 (1996).  
RN [4]  
RP SEQUENCE FROM N.A.

RC TISSUE=Lung;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feigold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T.H., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toehlyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: Forms a water-specific channel. Osmoreceptor which  
CC regulates body water balance and mediates water flow within the  
CC central nervous system.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=2;  
CC IsoID=P35087-1; Sequence=Displayed;  
CC Name=1;  
CC IsoID=P35087-2; Sequence=VSP\_003232;  
CC -!- TISSUE SPECIFICITY: Brain - muscle >> heart, kidney, lung, and  
CC trachea.  
CC -!- DOMAIN: Aquaporins contain two tandem repeats each containing  
CC three membrane-spanning domains and a pore-forming loop with the  
CC signature motif Aap-Pro-Ala (NPA).  
CC -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U34846; AAC52112.1; ALT INIT.  
CC EMBL: U34845; AAC50284.1; ALT INIT.  
CC EMBL: D63412; BAA09715.1; -.  
CC EMBL: U63622; AAB26957.1; -.  
CC EMBL: U63623; AAB26958.1; -.  
CC EMBL: BC022286; AAB22286.1; -.  
CC PIR: I39178; I39178.  
CC HSSP: P29972; IFOY.  
CC Genew: HGNC:637; AQP4.  
CC H-InvDB: HIX0014377; -.  
CC MIM: 600308; -.  
CC GO: GO:0005887; C:integral to plasma membrane; TAS.  
CC GO: GO:0005372; F:water transporter activity; TAS.  
CC GO: GO:0007588; P:excretion; TAS.  
CC GO: GO:0007399; P:neurogenesis; TAS.  
CC GO: GO:0005810; P:transport; TAS.  
CC InterPro: IPR000425; MIP.  
CC Pfam: PF00230; MIP; 1.  
CC PRINTS: PR00783; MINTRINSIC.  
CC ProDom: PD000295; MIP; 1.  
CC TIGRFAMs: TIGR00861; MIP; 1.  
CC PROSITE: PS00221; MIP; 1.  
CC Alternative splicing; Phosphorylation; Repeat; Transmembrane;  
KW Transport.  
FT DOMAIN 1 36 Cytoplasmic (Potential).  
FT TRANSMEM 37 57 Potential.

FT	DOMAIN	58	64	Extracellular (Potential).
FT	TRANSMEM	65	85	Potential.
FT	DOMAIN	86	115	Cytoplasmic (Potential).
FT	TRANSMEM	116	136	Potential.
FT	DOMAIN	137	155	Extracellular (Potential).
FT	TRANSMEM	156	176	Potential.
FT	DOMAIN	177	184	Cytoplasmic (Potential).
FT	TRANSMEM	185	205	Potential.
FT	DOMAIN	206	231	Extracellular (Potential).
FT	TRANSMEM	232	252	Potential.
FT	DOMAIN	253	323	Cytoplasmic (Potential).
FT	SITE	97	99	NPA 1.
FT	SITE	213	215	NPA 2.
FT	MOD_RES	285	285	Phosphoserine (By similarity).
FT	CARBOHYD	153	153	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	206	206	N-linked (GlcNAc. .) (Potential).
FT	VARSPLIC	1	22	Missing (in isoform 1). /FTid=VSP_003232.
FT	CONFLICT	246	246	G -> A (in Ref. 1).
FT	CONFLICT	287	288	VE -> AK (in Ref. 1).
FT	CONFLICT	296	296	P -> L (in Ref. 1).
SQ	SEQUENCE	323 AA; 34829 MW; 1A160C0F0DC11052 CRC64;		

Alignment Scores:  
Pred. No.: 1,41e-119 Length: 323  
Score: 1685.00 Matches: 323  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 77.26% Indels: 0  
DB: 1 Gaps: 0

US-10-723-180-1 (1-1152) x AQP4\_HUMAN (1-323)

QY	40	ATGAGTGCACAGCCACCAAGCGGTGGGGTGAAGTGTGGACCTTTGTGTACCAAGAGAG	99
Db	1	MetSerAspArgProThrAlaArgArgTrpGlyLysCysGlyProLeuCysThrArgGlu	20
QY	100	AACATCATGGTGGCTTCAAAAGGGGTCTGGACTCAAGCTTTCTGGAAGAGAGTCAACGG	159
Db	21	AsnileMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAla	40
QY	160	GAATTTCTGGCAGCTGTTATTTTCTCTCTCAGCGTGGGATCCACCATCACTGGGGT	219
Db	41	GlupheLeuAlaMetLeullePheValLeuLeuSerLeuGlySerThrIleAsnTrpGly	60
QY	220	GGAAACAGAAAAAGCCCTTTACCGGTGCACATGGTCTTCATCTCCCTTTGCTTGGACTCAGC	279
Db	61	GlyThrGluLysProLeuProValAspMetValLeulleSerLeuCysPheGlyLeuSer	80
QY	280	ATTGCAACCATGGTGCAGTGTCTTGGCCATATCAGCGTGGCCACATCAACCTGCAGTG	339
Db	81	IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal	100
QY	340	ACTGTGGCCATGGTGTGCACAGGAAGATCAGCATCGCAAGTCTGTCTTCTACATCCCA	399
Db	101	ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheThrIleAla	120
QY	400	GCCAGTGCCTGGGGGCCATCATTTGGAGCAGGAATCTCTTATCTGGTGCACCTCCCACT	459
Db	121	AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyLeuValThrProProSer	140
QY	460	GTGGTGGGAGCGCTGGAGTCCACATGGTCTTATCGGAATCTTACCGCTGGTGGTCTC	519
Db	141	ValValGlyGlyLeuGlyValThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeu	160
QY	520	CTGGTTGAGTTGATAATCACATTTCAATTTGGTGTGTACTATCTTTTGCAGCTGTGATCC	579
Db	161	LeuValGluLeuLeulleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer	180
QY	580	AAACGGACTGATGTCACCTGCTCAATAGCTTTAGCAATTTGATTTCTTGTTCGAATCGA	639
Db	181	LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGly	200

```

QY 640 CATTATTTGCAATCAATATATCTGCTGCAGCATCAATCCCGCCCGATCCTTTGGACCT 699
Db 201 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220
QY 700 GCAGTTTATCATCGGAATTTGGGAARACCATTCGATATATTTGGTTGGCCCATCATAGCA 759
Db 221 AlaValIleGlyLeuTyrGluAsnHsTrpIleTyrTrpValGlyProIleIleGly 240
QY 760 GCTGTCTCGCTGGTGGCCCTTTATGATGATGCTTCTCTGCTCCAGATGTTGAATTCAAACCT 819
Db 241 AlaValLeuAlaGlyLeuTyrGluTyrValPheCysProAspValGluPheLysArg 260
QY 820 CGTTTAAAGACCTTCAGCAAGCTGCCCGAGCAAAACAAAGGAAGCTACATGAGGTG 879
Db 261 ArgPheLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluVal 280
QY 880 GAGGACAACAGAGTCAGGTAGACGAGTACCTGATTTCTAAACCTGAGTGGTCAT 939
Db 281 GluAspAsnArgSerGlnValGluThrAspLeuIleLysProGlyValValHis 300
QY 940 GTGATTGACGTTGACCGGGGAGAGGAGGAAGGGGAAAGACCAATCTCGAGAGGTATTG 999
Db 301 ValIleAspValAspArgGlyGluGluLysLysGlyLysAspGlnSerGlyGluValLeu 320
QY 1000 TCTTCAGTA 1008
Db 321 SerSerVal 323

RESULT 2
AQP4_BOVIN
ID AQP4_BOVIN STANDARD; PRT; 323 AA.
AC 077750;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Aquaporin 4 (MCH4) (Mercurial-insensitive water channel) (MIWC).
GN Name=AQP4; Synonyms=AQP-4;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxId=9913;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=Holstein; TISSUE=Brain;
RX MEDLINE=20135596; PubMed=10673041; DOI=10.1016/S0167-4781(99)00194-3;
RA Sobue K., Yamamoto N., Yoneda K., Fujita K., Miura Y., Asai K.,
RA Tsuda T., Katsuya H., Kato T.;
RT "Molecular cloning of two bovine aquaporin-4 cDNA isoforms and their
RT expression in brain endothelial cells.";
RL Biochim. Biophys. Acta 1489:393-398(1999).
CC -!- FUNCTION: Forms a water-specific channel. Osmoreceptor which
CC regulates body water balance and mediates water flow within the
CC central nervous system.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=2; Synonyms=B;
CC IsoId=O77750-1; Sequence=Displayed;
CC Name=1; Synonyms=A;
CC IsoId=O77750-2; Sequence=VSP_003231;
CC -!- DOMAIN: Aquaporins contain two tandem repeats each containing
CC three membrane-spanning domains and a pore-forming loop with the
CC signature motif Asn-Pro-Ala (NPA).
CC -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL; AB015947; BAA36505.2; --
DR EMBL; AB012950; BAA33583.1; --
DR EMBL; AB028642; BAA89291.1; --
DR HSSP; P29972; 1FOY.
DR InterPro; IPR000425; MIP.
DR Pfam; PF00230; MIP; 1.
DR PRINTS; PR00783; MINTRINSICP.
DR ProDom; PD000295; MIP; 1.
DR TIGRFAMs; TIGR00861; MIP; 1.
DR PROSITE; PS00221; MIP; 1.
KW Alternative splicing; Phosphorylation; Repeat; Transmembrane;
KW Transport.
FT DOMAIN 1 36 Cytoplasmic (Potential).
FT TRANSMEM 37 57 Potential.
FT DOMAIN 58 64 Extracellular (Potential).
FT TRANSMEM 65 85 Potential.
FT DOMAIN 86 115 Cytoplasmic (Potential).
FT TRANSMEM 116 136 Potential.
FT DOMAIN 137 155 Extracellular (Potential).
FT TRANSMEM 156 176 Potential.
FT DOMAIN 177 184 Cytoplasmic (Potential).
FT TRANSMEM 185 205 Potential.
FT DOMAIN 206 231 Extracellular (Potential).
FT TRANSMEM 232 252 Potential.
FT DOMAIN 253 323 Cytoplasmic (Potential).
FT SITE 97 99 NPA 1.
FT SITE 213 215 NPA 2.
FT MOD_RES 285 285 Phosphoserine (By similarity).
FT CARBOHYD 153 153 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 206 206 N-linked (GlcNAc...) (Potential).
FT VARSPLIC 1 22 Missing (in isoform 1).
FT FTId=VSP_003231.
SQ SEQUENCE 323 AA; 34616 MW; E8E14507B6474D5F CRC64;
US-10-723-180-1 (1-1152) x AQP4_BOVIN (1-323)
Alignment Scores:
Pred. No.: 3,04e-115 Length: 323
Score: 1628.00 Matches: 311
Percent Similarity: 97.52% Conservative: 4
Best Local Similarity: 96.28% Mismatches: 8
Query Match: 74.64% Indels: 0
Db: 1 Gaps: 0
QY 40 ATGAGTGCAGACCCACAGCAGCGGTGGGTAGTGTGGACCTTTGTGTACCGAGAG 99
Db 1 MetSerAspArgProAlaAlaThrArgTrpGlyLysCysGlyProLeuCysThrArgGlu 20
QY 100 AACATCATGTTGGCTTTCAAAGGGGTCTGGACTCAAGCTTTCTGGAAGACGAGTCACAGCG 159
Db 21 SerIleMetValAlaPheLysGlyValTrpThrGlnThrPheTrpLysAlaValThrAla 40
QY 160 GAATTTCTGGCCATGCTTATTTTCTCTCAGCCTGGGATCCACCATCAACTGGGGT 219
Db 41 GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGly 60
QY 220 GGAACAGAAAAGCCTTTACCGGTGCACATGTTTCTCATCTCCCTTTGGTGGACTCAGC 279
Db 61 GlyAlaGluLysProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer 80
QY 280 ATTGCAACATGTTGCAGTGTCTTTGGCCATATCAGCGGTGGCCACATCAACCTCAGTG 339
Db 81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100
QY 340 ACTGTGGCCATGTTGTGCACCAGGAAGATCAGCATCGCCAAAGTCTGTCTTCTACATCGCA 399
Db 101 ThrValAlaMetValCysThrArgArgIleSerIleAlaLysSerValPheTyrIleAla 120
QY 400 GCCAGTGCCTGGGGGCCATCATTTGGAGCAGGAATCTCTATCTGGTGCACACCTCCAGT 459
Db 121 AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValThrProSer 140

```

```
QY 460 GTGGTGGGAGCCTGGAGTCCACATGTTTCATGGAATCTTACCGCTGCTCATGCTC 519
DB 141 ValValGlyGlyLeuGlyValThrValHisGlyAsnLeuSerAlaGlyHisGlyLeu 160
QY 520 CTGGTTGAGTTGATATCATCATTTCAATTTGGTGTCTTACTATCTTTGCCAGCTGATTC 579
DB 161 LeuValGluLeuLeuIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180
QY 580 AAACGGACATGTCCTACTGGCTCAATAGCTTTAGCAATTTGGATTTCTGTTCAATTGGA 639
DB 181 LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyIleSerValAlaIleGly 200
QY 640 CATTATTGTCATCAATATATCTGTCGACATGAATCCCGCCGATCCTTTGGACCT 699
DB 201 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220
QY 700 GCAGTTATCATGGAAATGGGAAACCAATTCATATATTTGGTTGGGCCCATCATAGGA 759
DB 221 AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly 240
QY 760 GCTGTCCTCGCTGGGCTTTATGAGTATGCTCTGTCGACATGTCCTGTCAGATGTTGAAATCA 819
DB 241 AlaValLeuAlaGlyLeuTyrGluTyrValPheCysProAspValGluLeuLysArg 260
QY 820 CTTTTTAAAGAACCTTCAGCAAGCTGCCAGCAACAAAGGAAGCTACATGGAGGTG 879
DB 261 ArgPheLysGluAlaPheSerLysAlaAlaGlnInThrLysGlySerTyrMetGluVal 280
QY 880 GAGGACAAAGGAGTCCAGTACGAGCGGATGACCTGATTTCTAAAACCTGGAGTGGTGCAT 939
DB 281 GluAspAsnArgSerGlnValGluThrAspAspLeuLeuLysProGlyValValHis 300
QY 940 GTGATTGAGCTTACCGGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 999
DB 301 ValIleAspIleAspArgGlyGluGluLysLysGlyLysAspProSerGlyGluValLeu 320
QY 1000 TCTTCAGTA 1008
DB 321 SerSerVal 323
```

## RESULT 3

```
Q866S4 PRELIMINARY; PRT; 323 AA.
AC Q866S4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Aquaporin 4A (Aquaporin-4 M1 isoform).
GN Name=AQP4A; Synonyms=AQP4;
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Cognlan M.W., Koukoulas I., Armugan A., Jeyaseelan K.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Bondar A.A., Zelenina M.N., Gunnarson E., Illarionova N.B., Aperia A.,
RA Zelenin S.M.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR EMBL; AV177612; AA021366.1; -;
DR EMBL; AY304480; AA074771.1; -;
DR HSSP; P29972; 1H61.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR GO; GO:0015288; F:porin activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
```

```
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000425; MIP.
DR Pfam; PF00230; MIP; 1.
DR PRINTS; PR01415; ANKYRIN.
DR PRINTS; PR00783; MINTRINSCP.
DR ProDom; PD000295; MIP; 1.
DR TIGRPFAMs; TIGR00861; MIP; 1.
DR PROSITE; PS00221; MIP; 1.
KW Porin; Transmembrane; Transport.
SQ SEQUENCE 323 AA; 34754 MW; D959292E21374948D CRC64;
```

```
Alignment Scores:
Pred. No.: 7,3e-115 Length: 323
Score: 1623.00 Matches: 310
Percent Similarity: 97.52% Conservative: 5
Best Local Similarity: 95.98% Mismatches: 8
Query Match: 74.42% Indels: 0
DB: 2 Gaps: 0
```

US-10-723-180-1 (1-1152) x Q866S4 (1-323)

```
QY 40 ATGAGTGCAGACCCACAGCAAGCGGTGGGTAAAGTGTGGACCTTTGTGTACCAAGAG 99
DB 1 MetSerAspArgProAlaAlaArgArgTrpGlyLysCysGlyProLeuCysThrArgGlu 20
QY 100 AACATCATGTGTGGCTTTCAAAGGGGCTCGGACTCAAGCTTTCTGGAAGAGCTCACAGG 159
DB 21 SerIleMetValAlaPheLysGlyValTrpThrGlnThrPheTrpLysAlaValThrAla 40
QY 160 GAATTTCTGGCCATGCTTATTTTGTCTCCTCAGCTGGGATCCACCATCAACTGGGGT 219
DB 41 GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGly 60
QY 220 GGAACAGAAAAAGCCCTTTACCGGTGCACATGTTCTTCATCTCCCTTTGTTGGACTCAGC 279
DB 61 GlyAlaGluLysProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer 80
QY 280 ATTGCAACCATGTGTGAGTCTTTGGCCATATCAGCGGTGGCCACATCAACCTGTCAGTG 339
DB 81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100
QY 340 ACTGTGGCCATGTGTGACCCAGGAAGATCAGCATCGCAAGTCTGTCTTACATCCCA 399
DB 101 ThrValAlaMetValCysThrArgGlySerIleAlaLysAlaValPheTyrIleAla 120
QY 400 GCCCAGTCCCTGGGGCCCATCATTTGGAGCAGGAATCTCTCTATCTGTGTGCACACCTCC 459
DB 121 AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValThrProProSer 140
QY 460 GTGTGGGAGCGCTGGGAGTCCACCATGTTTCATGGAATCTTACCGTGGTTCATGGTCTC 519
DB 141 ValValGlyGlyLeuGlyValThrValHisArgAsnLeuSerAlaGlyHisGlyLeu 160
QY 520 CTGGTTGAGTTGATAATCACATTTCAATTTGGTGTGTCTTACTATCTTGGCAGCTGTGATTC 579
DB 161 LeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180
QY 580 AAACGGACATGTCCTACTGGCTCAATAGCTTTAGCAATTTGGATTTCTGTTGCAATTGGA 639
DB 181 LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyIleSerValAlaIleGly 200
QY 640 CATTATTGTCATCAATATATCTGTCGACATGAATCCCGCCGATCCTTTGGACCT 699
DB 201 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220
QY 700 GCAGTTATCATGGAAATGGGAAACCAATTCATATATTTGGTTGGGCCCATCATAGGA 759
DB 221 AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly 240
QY 760 GCTGTCCTCGCTGGGCTTTATGAGTATGCTCTTCTGTCCAGATGTTGAAATCAACAGT 819
DB 241 AlaValLeuAlaGlyLeuTyrGluTyrValPheCysProAspValGluLeuLysArg 260
```



```

QY 820 CGTTTTAAAGACCTTCAGCAAGCTGCCAGCAACAAAGAAAGCTACATGAGGTG 879
D 261 AtgPheLysGluAlaPheSerLysAlaAlaGlnThrLysGlySerThrMetGluVal 280
QY 880 GAGGACAAAGAGTTCAGTAGAGCGATGACCTGATTTCTAAACCTGGAGTGGCAT 939
D 281 GluAspAsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGlyValValHis 300
QY 940 GTGATTGACGTTGACCGGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 999
D 301 ValLeuAspLeuAspArgGlyGluGluLysLysGlyAspProSerGlyGluValLeu 320
QY 1000 TCTTCAGTA 1008
D 321 SerSerVal 323

RESULT 4
AQP4_DIPME STANDARD; PRT; 324 AA.
AC Q923J4; Q923J5;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Aquaporin 4.
GN Name=AQP4;
OS Dipodomys merriami (Merriam's kangaroo rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Heteromyidae;
OC Dipodomys; Dipodomys.
OX NCBI_TaxId=94247;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RA Fang P.K., Huang Y., Walsberg G.E., Brown D., van Hoek A.N.;
RT "Molecular cloning and characterization of Merriam's kangaroo rat
RT aquaporin 4.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP TISSUE SPECIFICITY.
RX PubMed=11292621;
RA Huang Y., Tracy R., Walsberg G.E., Makkinje A., Fang P., Brown D.,
RA van Hoek A.N.;
RT "Absence of aquaporin-4 water channels from kidneys of the desert
RT rodent Dipodomys merriami merriami.";
RL Am. J. Physiol. 280:F794-F802(2001).
CC -!- FUNCTION: Forms a water-specific channel. Osmoreceptor which
CC regulates body water balance and mediates water flow within the
CC central nervous system (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=2;
CC IsoId=Q923J4-1; Sequences=Displayed;
CC Name=1;
CC IsoId=Q923J4-2; Sequences=VSP_010210;
CC -!- TISSUE SPECIFICITY: Not expressed in kidney. Detectable in gastric
CC parietal and brain astroglial cells. The absence of AQP4 in kidney
CC may be critical for the extreme urinary concentration that occurs
CC in this species (up to 5,000 mosmol/Kg H2O).
CC -!- DOMAIN: Aquaporins contain two tandem repeats each containing
CC three membrane-spanning domains and a pore-forming loop with the
CC signature motif Asn-Pro-Ala (NPA).
CC -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AY032857; AAK66823.1; -.

```

```

DR EMBL; AY032858; AAK66824.1; -.
DR HSP; P29972; 1FOY.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000425; MIP.
DR Pfam; PF00230; MIP; 1.
DR PRINTS; PRO1415; ANKYRIN.
DR PRINTS; PRO0783; MINTRINSICP.
DR ProDom; PD000295; MIP; 1.
DR TIGRFAMs; TIGR00861; MIP; 1.
DR PROSITE; PS00221; MIP; 1.
KW Alternative splicing; Phosphorylation; Repeat; Transmembrane;
KW Transport.
FT DOMAIN 1 36 Cytoplasmic (Potential).
FT TRANSMEM 37 57 Potential.
FT DOMAIN 58 63 Extracellular (Potential).
FT TRANSMEM 64 84 Potential.
FT DOMAIN 85 115 Cytoplasmic (Potential).
FT TRANSMEM 116 136 Potential.
FT DOMAIN 137 155 Extracellular (Potential).
FT TRANSMEM 156 176 Potential.
FT DOMAIN 177 184 Cytoplasmic (Potential).
FT TRANSMEM 185 205 Potential.
FT DOMAIN 206 231 Extracellular (Potential).
FT TRANSMEM 232 252 Potential.
FT DOMAIN 253 324 Cytoplasmic (Potential).
FT SITE 97 99 NPA 1.
FT SITE 213 215 NPA 2.
FT MOD_RES 285 Phosphoserine (By similarity).
FT CARBOHYD 153 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 206 N-linked (GlcNAc...) (Potential).
FT VARSPLIC 1 Missing (in isoform 1).
FT FTIG=VSP_010210.
SQ SEQUENCE 324 AA; 34939 MW; EA9CE3B17952CDD CRC64;

Alignment Scores:
Pred. No.: 8.25e-113 Length: 324
Score: 1596.00 Matches: 304
Percent Similarity: 96.87% Conservative: 5
Best Local Similarity: 95.30% Mismatches: 10
Query Match: 73.18% Indels: 0
DB: 1 Gaps: 0

US-10-723-180-1 (1-1152) x AQP4_DIPME (1-324)
QY 40 ATGAGTGCAGACCCACAGAGCGGTGGGTGAGTGGACCTTTGTGTACGAGAG 99
D 1 MetSerAspArgProAlaAlaArgProTrpGlyLysCysGlySerLeuCysArgGlu 20
QY 100 AACATCATGGTGGCTTTCAAAGGGGTCTGGACTCAAGCTTTCTGGAAGACGTCACAGCG 159
D 21 GluIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValThrAla 40
QY 160 GAATTTCTGGCCATGCTTATTTTCTCTCAGCTGGGATCCACCATCAACTGGGGT 219
D 41 GluPheLeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGly 60
QY 220 GGAACAGAAAAAGCCCTTACCGGTGCACATGGTTCATCTCCCTCTTGTTCGACTCAGC 279
D 61 GlyLysGluAsnProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer 80
QY 280 ATTGCAACCATGGTGCAGTGTCTTTCGCCATATCATCGGTGGCCACATCAACCTCGAGTG 339
D 81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100
QY 340 ACTGTGGCCATGGTGTGCACCAAGAGATCAGCATCGCAAGTGTCTTCTTACATCGCA 399
D 101 ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheThrIleAla 120
QY 400 GCCACGTCCCTGGGGCCCATCATTCGAGCAGAGAACTCTCTATCTGTGTACACCTCCAGT 459
D 121 AlaGlnCysLeuGlyAlaAlaIleGlyAlaGlyIleLeuTyLeuValThrProSer 140
QY 460 GTGGTGGGAGCGCTGGGAGTCACCATGGTTCATGGAATCTTACCGCTGGTTCATGCTCTC 519

```

```
Db 141 ValValGlyGlyLeuGlyValThrThrValHisGlyAsnLeuThraAlaGlyHisGlyLeu 160
QY 520 CTGGTGGAGTGAATACATTCATTAATGGTGGTTTACTATCTTCCAGCTGTGATCC 579
Db 161 LeuValGluLeuIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180
QY 580 AAACGGACTGATCTACTGGCTCAATAGCTTTAGCAATGGATTTCTGTTCATTTGA 639
Db 181 LysA-gThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGly 200
QY 640 CATTATTGGCAATCAATATATCTGTGCACATGAATCCCGCCGATCTTTGGACCT 699
Db 201 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220
QY 700 GCAGTTATCATCGGAATGGAAACCACTGGATATATTTGGTGGCCCATCATAGA 759
Db 221 AlaValIleMetGlyAsnTrpGluAsnHisTyrIleTyrTrpValGlyProIleIleGly 240
QY 760 GTGTCTCTCGTGGTGGCTTTATGAGTATGCTTCTGTCAGATGTTGAATTCAAAGT 819
Db 241 AlaValLeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluLeuLysArg 260
QY 820 CTTTTTAAAGAGCTTTCAGCAAGCTGCCAGCAAAAGAGAGGAGGCTACATGAGGTG 879
Db 261 ArgPheLeuGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluVal 280
QY 880 GAGGACACAGAGTCAAGTACAGACGAGTACCTGATTTCTAAACCTGGAGTGGTGCAT 939
Db 281 GluAspAsnArgSerGlnValGluThrGluAspLeuLeuLysProGlyLeuValHis 300
QY 940 GTGATTCAGTTCAGCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGA 996
Db 301 ValIleAspIleAspArgGlyAspGluLysGlyLysGlyLysAspProSerGlyGluIle 319

RESULT 5
AQP4_RAT
ID AQP4_RAT STANDARD; PRT; 323 AA.
AC F47863;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Aquaporin 4 (WCH4) (Mercurial-insensitive water channel) (MIWC).
GN Name=Aqp4;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1] _TaxID=10116;
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF HIS-201.
RC TISSUE=Brain;
RX MEDLINE=95108097; PubMed=7528931;
RA Jung J.S., Bhat R.V., Preston G.M., Guggino W.B., Baraban J.M.,
RA Agre P.;
RT "Molecular characterization of an aquaporin cDNA from brain: candidate
RT osmoreceptor and regulator of water balance.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:13052-13056(1994).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Chen D., Chen J., Jing K., Simon R.P., Graham S.H.;
RT "Isolation of an aquaporin-4 water channel (AQP4) gene induced
RT following cerebral ischemia from the rat brain using modified
RT subtractive hybridization and differential screening.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBSJ databases.
[3]
RP SEQUENCE OF 23-323 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94164885; PubMed=7509789;
RA Hasegawa H., Ma T., Skach W., Matthay M.A., Verkman A.S.;
RT "Molecular cloning of a mercurial-insensitive water channel expressed
RT in selected water-transporting tissues.";
RL J. Biol. Chem. 269:5497-5500(1994).
```

```
[4]
RN PHOSPHORYLATION SITE SER-285.
RP PubMed=12692561; DOI=10.1038/nbt819;
RX Wu C.C., MacCoss M.J., Howell K.E., Yates J.R. III;
RT "A method for the comprehensive proteomic analysis of membrane
RT proteins.";
RL Nat. Biotechnol. 21:532-538(2003).
CC -|- FUNCTION: Forms a water-specific channel. Osmoreceptor which
CC regulates body water balance and mediates water flow within the
CC central nervous system. It is expressed predominantly in the
CC ependymal cell lining the aqueductal system and over the space of
CC the brain in contact with the subarachnoid space, as cerebrospinal
CC fluid fills these structures it may facilitate water balance
CC between brain parenchyma and the fluid compartment. In the plasma
CC membranes of the neurons of the paraventricular and supraoptic
CC nuclei, it may mediate rapid changes in cell volume in response to
CC local shifts in extracellular osmolality.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P47863-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P47863-2; Sequence=VSP_003235;
CC -|- TISSUE SPECIFICITY: Abundant in mature brain but only weakly
CC detectable in eye, kidney, intestine, and lung.
CC -|- DOMAIN: Aquaporins contain two tandem repeats each containing
CC three membrane-spanning domains and a pore-forming loop with the
CC signature motif Asn-Pro-Ala (NPA).
CC -|- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
CC -|- CAUTION: It is uncertain whether Met-1 or Met-23 is the initiator.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; U14007; AAC52152.1; -.
EMBL; AF144082; AAD37965.1; -.
EMBL; L27588; AAA17730.1; -.
PIR; I59283; I59283.
HSSP; P29972; IFOY.
RGD; 2143; Aqp4.
InterPro; IPR000425; MIP.
Pfam; PF00230; MIP; 1.
PRINTS; PR00783; MINTRINSCP.
PRODOM; PD000295; MIP; 1.
TIGRFAMs; TIGR00861; MIP; 1.
PROSITE; PS00221; MIP; 1.
Alternative splicing; Phosphorylation; Repeat; Transmembrane;
Transport.
DOMAIN 1 36 Cytoplasmic (Potential).
TRANSMEM 37 57 Potential.
DOMAIN 58 64 Extracellular (Potential).
TRANSMEM 65 85 Potential.
DOMAIN 86 115 Cytoplasmic (Potential).
TRANSMEM 116 136 Potential.
DOMAIN 137 155 Extracellular (Potential).
TRANSMEM 156 176 Potential.
DOMAIN 177 184 Cytoplasmic (Potential).
TRANSMEM 185 205 Potential.
DOMAIN 206 231 Extracellular (Potential).
TRANSMEM 232 252 Potential.
DOMAIN 253 323 Cytoplasmic (Potential).
SITE 97 NPA 1.
SITE 213 NPA 2.
MOD_RES 285 Phosphoserine.
CARBOHYD 153 N-linked (GlcNAc..) (Potential).
VARSPIC 150 Missing (in isoform Short).
FT /FTid=VSP_003235.
FT
```

940 CTCATTGACCTTGATCCGGAGAGAGAGAGAGAGCGGGAAGCAACCATTCGTGAGAGCTATTG 99

301 ValIIIeAspIIleAspArgGIyAspGIuIysLeysGLySerpSerGIyGluValLeu 320

1000 TCTTCAGTA 1008

321 SerSerVal 323

RESULT 6

AQP4 MOUSE

ID AQP4 MOUSE STANDARD; PRT: 323 AA.

AC P508B; P97818; Q61131; Q61132; Q8VHE4; Q8VHES; Q9EQI3;

DT 01-OCT-1996 (Rel. 34, Created)

DT 05-JUL-2004 (Rel. 44, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Aquaporin 4 (WCH4) (Mercurial-insensitive water channel) (MIWC).

GN Name:Aqp4;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCB\_I\_Taxid=10090;

[1]

RN SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).

RP STRAIN=C57BL/6; TISSUE=Brain;

RX MEDLINE=96299634; PubMed=8660998; DOI=10.1006/geno.1996.0214;

RA Ma T., Yang B., Verkman A.S.;

RT "Gene structure, cDNA cloning, and expression of a mouse mercurial-

RL insensitive water channel.";

RL Genomics 33:382-388(1996).

[2]

RN SEQUENCE FROM N.A. (ISOFORM 2).

RP STRAIN=C57BL/6J; TISSUE=Cerebellum;

RX MEDLINE=97728826; PubMed=9143504; DOI=10.1006/geno.1997.4641;

RA Turtzo L.C., Lee M.D., Lu M., Smith B.L., Copeland N.G., Gilbert D.J.,

RA Jenkins N.A., Agre P.;

RT "Cloning and chromosomal localization of mouse aquaporin 4: exclusion

RL of a candidate mutant phenotype, ataxia.";

RL Genomics 41:267-270(1997).

[3]

RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANT ARG-4.

RP MEDLINE=20419017; PubMed=10960499;

RX Zelenin S., Gunnarson E., Alilikina T., Bondar A., Aperia A.;

RT "Identification of a new form of AQP4 mRNA that is developmentally

RL expressed in mouse brain".

RL Pediatr. Res. 48:335-339(2000).

[4]

RN SEQUENCE FROM N.A. (ISOFORM 1).

TISSUE=Colon;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Pahay J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Wadan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RL and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).

CC -!- FUNCTION: Forms a water-specific channel. Osmoreceptor which

CC regulates body water balance and mediates water flow within the

CC central nervous system.

CC CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC CC -!- ALTERNATIVE PRODUCTS;  
 CC CC Event=Alternative splicing; Named isoforms=3;  
 CC CC Name=2; Synonyms=MIWC2, AQP4-M1;  
 CC CC IsoId=P55088-1; SequencesDisplayed;  
 CC CC Name=1; Synonyms=MIWC1, AQP4-M21;  
 CC CC IsoId=P55088-2; Sequence=VSP\_003233;  
 CC CC Name=3; Synonyms=MIWC3;  
 CC CC IsoId=P55088-3; Sequences=VSP\_003234;  
 CC CC -!- DOMAIN: Aquaporins contain two tandem repeats each containing  
 CC CC three membrane-spanning domains and a pore-forming loop with the  
 CC CC signature motif Asn-Pro-Ala (NPA).  
 CC CC -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.  
 CC CC -----  
 CC CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC CC the European Bioinformatics Institute. There are no restrictions on its  
 CC CC use by non-profit institutions as long as its content is in no way  
 CC CC modified and this statement is not removed. Usage by and for commercial  
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC CC -----  
 CC DR EMBL; U48398; AAB41569.1; -;  
 CC DR EMBL; U48397; AAB41568.1; -;  
 CC DR EMBL; U33012; AAB4923.1; -;  
 CC DR EMBL; U48400; AAB41571.1; -;  
 CC DR EMBL; U48399; AAB41570.1; -;  
 CC DR EMBL; U88623; AAC53155.1; -;  
 CC DR EMBL; AF469168; AAL73545.1; -;  
 CC DR EMBL; AF469169; AAL73546.1; -;  
 CC DR EMBL; AF219992; AAG44243.2; -;  
 CC DR EMBL; BC024526; AAB424526.1; -;  
 CC DR HSP; F29972; IFQ1.  
 CC DR MGD; MGI:107387; Aqp4.  
 CC DR GO; GO:0016021; C:integral to membrane; NAS.  
 CC DR GO; GO:0015250; P:water channel activity; IDA.  
 CC DR GO; GO:0030104; P:water homeostasis; NAS.  
 CC DR InterPro; IPR000425; MIP.  
 CC DR Pfam; PF00230; MIP; 1.  
 CC DR PRINTS; PR00783; MINTRINSICP.  
 CC DR ProDom; PD00295; MIP; 1.  
 CC DR TIGRFAMs; TIGR00861; MIP; 1.  
 CC DR PROSITE; PS00221; MIP; 1.  
 CC KW Alternative splicing; Phosphorylation; Repeat; Transmembrane;  
 CC KW Transport.  
 CC FT DOMAIN 1 36 Cytoplasmic (Potential).  
 CC FT TRANSMEM 37 58 Potential.  
 CC FT DOMAIN 59 64 Extracellular (Potential).  
 CC FT TRANSMEM 65 85 Potential.  
 CC FT DOMAIN 86 115 Cytoplasmic (Potential).  
 CC FT TRANSMEM 116 136 Potential.  
 CC FT DOMAIN 137 155 Extracellular (Potential).  
 CC FT TRANSMEM 156 176 Potential.  
 CC FT DOMAIN 177 184 Cytoplasmic (Potential).  
 CC FT TRANSMEM 185 205 Potential.  
 CC FT DOMAIN 206 231 Extracellular (Potential).  
 CC FT TRANSMEM 232 252 Potential.  
 CC FT DOMAIN 253 323 Cytoplasmic (Potential).  
 CC FT SITE 97 99 NPA 1.  
 CC FT SITE 213 215 NPA 2.  
 CC FT MOD\_RES 285 285 Phosphoserine (By similarity).  
 CC FT CARBOHYD 153 153 N-linked (GlcNAc...) (Potential).  
 CC FT CARBOHYD 206 206 N-linked (GlcNAc...) (Potential).  
 CC FT VARSPLIC 1 22 Missing (in isoform 1).  
 CC FT FTID=VSP\_003233.  
 CC FT VARSPLIC 1 11 MSDGAAARRWG -> MVHFGFCVFFFLISLSLWASEDST  
 CC FT CNST:PLCHLATLDDC (in isoform 3).  
 CC FT FTID=VSP\_003234.  
 CC FT G -> R.  
 CC FT VARIANT 4 4 A -> S (in Ref. 1; AAB41571).  
 CC FT CONFLICT 37 37 Missing (in Ref. 1).  
 CC FT CONFLICT 51 51 F -> L (in Ref. 1).  
 CC FT CONFLICT 88 88 I -> V (in Ref. 1).  
 CC FT CONFLICT 174 174

FT CONFLICT 313 313 K -> R (in Ref. 1).  
 SQ SEQUENCE 323 AA; 34436 MW; 9AC638A3C3F412E2 CRC64;  
 Alignment Scores:  
 Pred. No.: 1.32e-110 Length: 323  
 Score: 1567.00 Matches: 300  
 Percent Similarity: 95.36% Conservative: 8  
 Best Local Similarity: 92.88% Mismatches: 15  
 Query Match: 71.85% Indels: 0  
 DB: 1 Gaps: 0  
 US-10-723-180-1 (1-1152) x AQP4\_MOUSE (1-323)  
 QY 40 ATGAGTGCACAGCCACAGCAAGCGGTGGGTGAAGTGTGGACCTTTGTGTACACGAGAG 99  
 Db 1 MetSerAspGlyAlaAlaAlaArgTrpGlyLysCysGlyHisSerCysSerArgGlu 20  
 QY 100 AACATCATGTGGCTTTCAAAAGGGTCTGGACTCAAGCTTCTGGAAGAGCAGTACACGG 159  
 Db 21 SerIleMetValAlaPheLysGlyValTrpThrGlnAlaPheTrpLysAlaValSerAla 40  
 QY 160 GAATTCTGGCCATGCTTATTTTCTCTCAGCCTGGGATCCACCATCAACTGGGGT 219  
 Db 41 GluPheLeuAlaThrLeuIlePheValLeuLeuGlyValGlySerThrIleAsnTrpGly 60  
 QY 220 GGAACAGAAAAGCCTTTACCGGTGCAGATGTTCTCATCTCCCTTGGTGGACTCAGC 279  
 Db 61 GlySerGluAenProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSer 80  
 QY 280 ATTGCAACCATGTGTGCTGTTGGCCATATCAGCGTGGCCATCAACCCCTGACGTG 339  
 Db 81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100  
 QY 340 ACTGTGGCCATGTGTGCACAGGAAGATCAGCATCGCAAGTCTGTCTTCTACATCGCA 399  
 Db 101 ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIleIle 120  
 QY 400 GCCAGTGCCTGGGGGCCATCATTTGGAGCAGGAATCTCTATCTGGTGCACACCTCCAGT 459  
 Db 121 AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValTrpProSer 140  
 QY 460 GTGTGGGAGCCTGGGAGTCACCATGTTTCATGGAATCTTACCGCTGGTGCATGCTCTC 519  
 Db 141 ValValGlyGlyLeuGlyValThrValHisGlyAsnLeuThrAlaGlyHisGlyLeu 160  
 QY 520 CTGTTGAGTTGATAATCACATTTCAATTTGGTGTGTACTATCTTTGCAGCTGTGATTC 579  
 Db 161 LeuValGluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 180  
 QY 580 AAACGGACTGATGCACCTGGCTCAATAGCTTTAGCAATTTGGATTTCTGTTGCAATTGGA 639  
 Db 181 LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGly 200  
 QY 640 CATTTATTGCAATCAATTAATCTGGTGCCAGCATGAATCCGCCGCGATCCTTTGGACCT 699  
 Db 201 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 220  
 QY 700 GCAGTTATCATGGGAAATTTGGGAAACCATTTGGATATATTGGGTGGGCCCATCATAGA 759  
 Db 221 AlaValIleMetGlyAsnTrpAlaAsnHisTrpIleTyrTrpValGlyProIleMetGly 240  
 QY 760 GCTGCTCTCGTGGTGGCTTTATGAGTATGTTCTCTCTCCAGATGTTGAATTCAAAGT 819  
 Db 241 AlaValLeuAlaGlyAlaLeuTyrGluTyrValPheCysProAspValGluLeuLysArg 260  
 QY 820 CGTTTAAAGAGCCTTCAGCAAGAGCTCCAGCAAAACAAAAGGAAGCTATCATGGAGGTG 879  
 Db 261 ArgLeuLysGluAlaPheSerIleAlaAlaGlnThrLysGlySerTyrMetGluVal 280  
 QY 880 GAGGACAAACAGAGTCAAGTGTAGACCGGATCACTGATTCTTAAACCTGGAGTGGTGCAT 939  
 Db 281 GluAspAenArgSerGlnValGluThrGluAspLeuIleLeuLysProGlyValValHis 300

```

QY 940 GTGATTGACGTTGACCGGGGAGAGAGAAAGGGGAAAGACCAATCTGGAGGATTTG 999
Db 301 ValIleAspIleAspArgGlyGluGluLysGlyLysAspSerSerGlyGluValLeu 320
QY 1000 TCTTCAGTA 1008
Db 321 SerSerVal 323

RESULT 7
Q6XVT6
ID Q6XVT6 PRELIMINARY; PRT; 301 AA.
AC Q6XVT6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Aquaporin 4 M23 isoform.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]_TaxID=9940;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Bondar A.A., Zelenin S., Gunnarson E., Aperia A.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR EMBL; AY208465; AAO38843.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR GO; GO:0015288; F:porin activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008810; P:transport; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000425; MIP.
DR Pfam; PF00230; MIP; 1.
DR PRINTS; PR01415; ANKYRIN.
DR PRINTS; PR00783; MINTRINSICP.
DR ProDom; PD000295; MIP; 1.
DR TIGRfam; TIGR00861; MIP; 1.
DR PROSITE; PS00221; MIP; 1.
KW Porin; Transmembrane; Transport.
SQ SEQUENCE 301 AA; 32282 MW; 36417087A5B2APE9 CRC64;

Alignment Scores:
Pred. No.: 6.82e-106 Length: 301
Score: 1505.00 Matches: 290
Percent Similarity: 97.67% Conservative: 4
Best Local Similarity: 96.35% Mismatches: 7
Query Match: 69.01% Indels: 0
DB: 2 Gaps: 0

US-10-723-180-1 (1-1152) x Q6XVT6 (1-301)

QY 106 ATGGTGGCTTCAAGGGGTCTGGACTCAAGCTTTCTGGAAAGCAGTCACAGCGAATTT 165
Db 1 MetValAlaPheLysGlyValIleThrPheThrPheLysAlaValThrAlaGluPhe 20
QY 166 CTGGCCATGCTTATTTTCTCTCTCAGCCTGGGATCCACCATCAACTGGGGTGAACA 225
Db 21 LeuAlaMetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGlyAla 40
QY 226 GAAAGCCTTTACCGGTGACATGTTCTCATCTCCCTTTGCTTGGACTCAGCATGCA 285
Db 41 GluLysProLeuProValAspMetValLeuIleSerLeuCysPheGlyLeuSerIleAla 60
QY 286 ACCATGGTCAGTCTTTCGCCATATCAGGGTGGCCACATCAACCTGCAGTGTG 345
Db 61 ThrMetValGlnCysPheGlyHisIleSerGlyHisIleAsnProAlaValThrVal 80
QY 346 GCCATGGTGTGCACAGGAAGATCAGCATCGCGCAAGTCTGTCTTACATCGAGCCAG 405

```

```

Db 81 AlaMetValCysThrArgIleSerIleAlaLysAlaValPheIleAlaGln 100
QY 406 TGCCCTGGGGCCCATATTGGACGAGGAATCCTCTATCTGGTCACACCTCCAGTGGTG 465
Db 101 CysLeuGlyAlaIleGlyAlaGlyIleLeuIleValThrProProSerValVal 120
QY 466 GGAGCCCTGGGAGTCACCATGGTTCATGGAAATCTTACCCTGGTGCATGCTCTCTGGTT 525
Db 121 GlyGlyLeuGlyValThrValHisArgAsnLeuSerAlaGlyHisGlyLeuVal 140
QY 526 GAGTTGATAATCACATTTCAATTGGTGTATTACTATCTTTGGCAGCTGTGATCCAAACGG 585
Db 141 GluLeuIleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSerLysArg 160
QY 586 ACTGATGTCACGTGCTCAATACGCTTTAGCAATTTGATTTCTGTTCGAATTCGACATTTA 645
Db 161 ThrAspValThrGlySerIleAlaLeuAlaIleGlyIleSerValAlaIleGlyHisLeu 180
QY 646 TTTGCAATCAATTATATCTGTCGCCAGCATGAATCCCGCCGATCCTTTGGAGCTGCAGTT 705
Db 181 PheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyProAlaVal 200
QY 706 ATCATGGGAAATTTGGGAAACCAATGGATATATGGTGGTGGCCCAATCATAGAGCTGTC 765
Db 201 IleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGlyAlaVal 220
QY 766 CTCGCTGGTGGCTTTATGAGTATGCTCTCTGTCAGATGTTGAATTCAAAGCTCGTTT 825
Db 221 LeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluLeuLysArgPhe 240
QY 826 AAGAAGCCTTCAGCAAGCTCCCGACGAAACAAAAGGAAGCTACATGAGGTGGAGGAC 885
Db 241 LysGluAlaPheSerLysAlaAlaGlnThrLysGlySerTyrMetGluValGluAsp 260
QY 886 AACAGGAGTCAGGTAGAGACGATGACCTGATTCTTAAACCTGGAGTGGTGCATGTGATT 945
Db 261 AsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGlyValValHisValIle 280
QY 946 GACGTTGACCGGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1005
Db 281 AspIleAspArgGlyGluGluLysGlyLysAspProSerGlyGluValLeuSerSer 300
QY 1006 GTA 1008
Db 301 Val 301

RESULT 8
Q6XVT6
ID Q6XVT6 PRELIMINARY; PRT; 335 AA.
AC Q6XVT6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Aquaporin-4.
GN Name=AQP4;
OS Coturnix coturnix (Common quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=9091;
RN [1]_TaxID=9091;
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Cui Y., Yang Y., Bolan J., Wang W., Nishimura H., Fan Z.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR EMBL; AF45730; AAL73511.1; -.
DR HSP; P29972; I1H5.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR GO; GO:0015288; F:porin activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.

```

```
DR GO: 0006810; P: transport; IEA.
DR Pfam: PF00230; MIP; 1.
DR PRINTS: PRO1415; ANKYRIN.
DR ProDom: PD00783; MINTRINSICP.
DR TIGRFAMs: TIGR00861; MIP; 1.
DR PROSITE: PS00221; MIP; 1.
KW Porin, Transmembrane, Transport.
SQ SEQUENCE 335 AA; 36164 MW; F942D4DA6F4D8DA CRC64;

Alignment Scores:
Pred. No.: 6.37e-103 Length: 335
Score: 1466.00 Matches: 276
Percent Similarity: 92.79% Conservative: 20
Best Local Similarity: 86.52% Mismatches: 23
Query Match: 67.22% Indels: 0
DB: 2 Gaps: 0

US-10-723-180-1 (1-1152) x Q80VB8 (1-335)

QY 52 CCCACAGCAAGCGGTGGGTAAAGTGTGGACCTTTGTGTACAGAGAGAACATCATGTGTG 111
DB 17 ProProAlaAAsSerSerLysCysGlyArgLeuLysCysGluAsnIleMetVal 36
QY 112 GCTTTCAAAGGGTCTGGACTCAAGCTTTCTGGAAGCAGTCACAGCGGAATTTCTGGCC 171
DB 37 AlaPheLysGlyValTrpThrGlnProPheTrpLysAlaValSerAlaGluPheLeuAla 56
QY 172 ATGCTTATTTTGTCTCTCAGCCTGGGATCCACCATCAACTGGGGTGGAAACGAAAG 231
DB 57 MetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGlySerGluLys 76
QY 232 CTTTACCGGTGCACATGTTCTCATCTCCCTTGTGCTTTGGAGTCAGCATTCGACCCATG 291
DB 77 ProLeuProValAspMetValLeuIleSerLeuLysPheGlyLeuSerIleAlaThrMet 96
QY 292 GNGCAGTCTTTGGCCATATCAGCGTGGCCACATCAACCTGCAGTCACTGCGCATG 351
DB 97 ValGlnCysPheGlyHisIleSerGlyHisIleAsnProAlaValThrValAlaMet 116
QY 352 GTGTGCACAGGAAGATCAGCATGCCAAGTCTGTCTTCTACATCGCAGCCAGTGCCTG 411
DB 117 ValCysThrArgLysIleSerLeuAlaLysSerValPheTrpIleLeuAlaGlnCysLeu 136
QY 412 GGGGCCATCTTTGGAGCAGGAATCTCTATCTGTGTACACCTCCAGTGTGTGGAGGC 471
DB 137 GlyAlaIleValGlyAlaGlyIleLeuTrpIleLeuIleThrProSerValValGlyGly 156
QY 472 CTGGGAGTCACCATGTTTCATCGAATCTTACCGCTGGTCAATGCTCTCGTTCAGTTG 531
DB 157 LeuGlyValThrAlaValHisGlyAspLeuSerAlaGlyHisGlyLeuLeuValGluLeu 176
QY 532 ATAATCACATTTCAATTGGTGTACTATCTTTTGGCCAGTGTGATTCCAAACGGACTGAT 591
DB 177 IleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSerLysArgSerAsp 196
QY 592 GTCACGTGCTCAATAGCTTTAGCAATTTGATTTCTTTGCAATTTGCAATTTATTTGCA 651
DB 197 ValThrGlySerValAlaLeuAlaIleGlyPheSerValAlaIleGlyHisLeuPheAla 216
QY 652 ATCAATTATPACTGTGGCCAGCATGAATCCCGCCGATCCTTTGGACCTGACGTTATCATG 711
DB 217 IleAsnTrpThrGlyAlaSerMetAsnProAlaLysSerPheGlyProAlaValIleMet 236
QY 712 GGAATTTGGGAAACCATTTGGATATATTTGGGTGGGCCCATCATAGGAGCTCTCTCGCT 771
DB 237 GlyLysTrpGluAsnGlnTrpValTrpValGlyProIleIleGlyAlaValLeuAla 256
QY 772 GGTGGCTTTATGATATGTTCTTGTCCAGATGTTGAATTCAAAGCTGTTTAAAGAA 831
DB 257 GlyAlaLeuTrpGluTrpValTrpCysProAspValGluLeuLysArgArgPheLysAsp 276
QY 832 GCCTTCAGCAAGCTGCCAGCAAAACAAAGAAAGCTACATCGAGGTGGAGGACACAGG 891

DR 277 ValPheSerLysThrSerGlnProSerLysGlyLysTrpIleGluValAspAspThrArg 296
QY 892 AGTCAGGTAGAGACGATGACCTGATTTCTTAAACCTGGAGTGGTGCATGTGATTCACGTT 951
DB 297 SerHisValGluThrAspAspLeuIleLeuLysProGlyIleValHisValIleAspIle 316
QY 952 GACCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1008
DB 317 AspArgSerGluAspLysGlyArgAspProSerSerGluValLeuSerSerVal 335

RESULT 9
Q65YQ3 PRELIMINARY; PRT; 335 AA.
AC Q65YQ3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Aquaporin 4.
GN Name=AQP4;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Saito N., Ikegami H., Shimada K.;
RT "cDNA cloning and mRNA expression of aquaporin 4 (AQP 4) in the
RL chicken (Gallus domesticus).";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB190358; BAD46731.1; -.
KW Porin.
SQ SEQUENCE 335 AA; 36258 MW; 66AE0B47CE75F6E5 CRC64;

Alignment Scores:
Pred. No.: 1.25e-101 Length: 335
Score: 1449.00 Matches: 273
Percent Similarity: 91.85% Conservative: 20
Best Local Similarity: 85.58% Mismatches: 26
Query Match: 66.44% Indels: 0
DB: 2 Gaps: 0

US-10-723-180-1 (1-1152) x Q65YQ3 (1-335)

QY 52 CCCACAGCAAGCGGTGGGTAAAGTGTGGACCTTTGTGTACAGAGAGAACATCATGTGTG 111
DB 17 ProProAlaAAsSerSerLysCysAlaArgLeuLysCysGlySerIleMetVal 36
QY 112 GCTTTCAAAGGGTCTGGACTCAAGCTTTCTGGAAGCAGTCACAGCGGAATTTCTGGCC 171
DB 37 AlaPheLysGlyValTrpThrHisProPheTrpLysAlaValSerAlaGluPheLeuVal 56
QY 172 ATGCTTATTTTGTCTCTCAGCCTGGATCCACCATCAACTGGGGTGGAAACGAAAG 231
DB 57 MetLeuIlePheValLeuLeuSerLeuGlySerThrIleAsnTrpGlySerGluLys 76
QY 232 CTTTACCGGTGCACATGTTCTCATCTCCCTTGTGCTTTGGAGTCAGCATTCGACCCATG 291
DB 77 ProLeuProValAspMetValLeuIleSerLeuLysPheGlyLeuSerIleAlaThrMet 96
QY 292 GTGCAGTCTTTGGCCATATCAGCGTGGCCACATCAACCTGCAGTCACTGTGGCCATG 351
DB 97 ValGlnCysPheGlyHisIleSerGlyHisIleAsnProAlaValThrValAlaMet 116
QY 352 GTGTGCACAGAGATCAGCATGCCAAGTCTGTCTTCTACATCGCAGCCAGTGCCTG 411
DB 117 ValCysThrArgLysIleSerLeuAlaLysSerValPheTrpIleLeuAlaGlnCysLeu 136
QY 412 GGGGCCATCTTTGGAGCAGGAATCTCTATCTGTGTACACCTCCAGTGTGTGGAGGC 471
DB 137 GlyAlaIleValGlyAlaGlyIleLeuTrpIleLeuIleThrProSerValValGlyGly 156
```



```
QY 472 CTGGGAGTCACCATGGTTCATGGAATCTTACCGCTGGTTCATGCTCTCTGCTTCAAGTTG 531
Db 157 LeuGlyValThrAlaValHisGlyAspLeuSerAlaGlyHisGlyLeuValGluLeu 176
QY 532 ATAATCACATTTCAATTGGTGTCTACTATCTTTGCCAGCTGTGATTCACAAACGGACTGAT 591
Db 177 IleIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSerIysArgSerAsp 196
QY 592 GTCACCTGGCTCAATAGCTTTAGCAATTTGGATTTCTTGTTGCAATTTGGACATTTATTGCA 651
Db 197 ValThrGlySerValAlaLeuAlaIleGlyPheSerValAlaIleGlyHisLeuPheAla 216
QY 652 ATCAATTATCTGCTGCCAGCATGAATCCGCCGCGATCTTTGGACCTGCAGTTATCATG 711
Db 217 IleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyProAlaValIleMet 236
QY 712 GCAAAATTTGGAACACCATGGATATATTCGGTTGGCCCATCATAGGAGCTGCTCTCGCT 771
Db 237 GlyIysTrpGluAsnGlnTrpValTyrTrpValGlyProIleIleGlyAlaValLeuAla 256
QY 772 GTCGCTTTATCAGTATGCTCTCTCCAGATGTTTGAATTCAAACGTCGTTTAAAGAA 831
Db 257 GlyAlaLeuTyrGlnTyrValTyrCysProAspValGluLeuLysArgGlyPheLysAsp 276
QY 832 GCTTCAGCAAAAGCTCCCGAGCAACAAAGGAAGCTACATGAGGTGGAGGACACAGG 891
Db 277 ValPheSerLysAlaThrGlnProSerLysGlyLysTyrIleGluValAspThrArg 296
QY 892 AGTCAGGTAGACGATGACCTGATCTTAAACCTGGAGTGGTGCATGCTGATTCAGCTT 951
Db 297 SerHisValGluThrAspAspLeuIleLeuLysProGlyIleValHisValIleAspIle 316
QY 952 GACCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1008
Db 317 AspArgSerGluAspLysGlyArgAspProSerSerGluValLeuSerSerVal 335

RESULT 10
Q8BR89 PRELIMINARY; PRT; 224 AA.
AC Q8BR89;
DT 01-WAR-2003 (TrEMBLrel. 23, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length
DE enriched library, clone:B23010SH10 product:aquaporin 4, full insert
DE sequence. (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RN FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RN FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=2549374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RN FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao K., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RN FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Horii F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kanokawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito K., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR EMBL; AK045357; BAC32325.1; -.
DR HSSP; P29972; 1HE1.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019867; C: outer membrane; IEA.
DR GO; GO:0015288; F: porin activity; IEA.
DR GO; GO:0005215; P: transporter activity; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000425; MIP.
DR Pfam; PF00230; MIP; 1.
DR PRINTS; PR01415; ANKYRIN.
DR PRINTS; PR00783; MINTRINSIC.
DR ProDom; PD000295; MIP; 1.
KW Porin; Transmembrane; transport.
FT NON TER
SQ SEQUENCE 224 AA; 23938 MW; 6076CD56ACEED029 CRC64;

Alignment Scores:
Pred. No.: 2,46e-75 Length: 224
Score: 1103.00 Matches: 214
Percent Similarity: 96.88% Conservative: 3
Best local Similarity: 95.54% Mismatches: 7
Query Match: 50.57% Indels: 0
DB: Gaps: 0

US-10-723-180-1 (1-1152) x Q8BR89 (1-224)
QY 337 GTGACTGTGGCCATGCTGTGCACGAGGAGATCAGCATCGCCAGCTGTCTTCTTACATC 396
Db 1 ValThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheTyrIle 20
QY 397 GCAGCCCAAGTGCCTGGGGCCCATCATTCGAGCAGGAATCTCTATCTGGTACACCTCCC 456
Db 21 IleAlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyrLeuValThrPro 40
```



```
QY 457 AGTGTGGTGGAGGCTGGAGTACCACATGGTTTCATGGAAATCTTACCGCTGGTCATGGT 516
Db |||||
QY 517 CTCCTGGTTGAGTGTGATATACATTCATTAATGGTGTCTTACTATCTTTGCCAGCTGTGAT 576
Db |||||
QY 61 LeuLeuValGluLeuLeuLeuPheGlnLeuValPheThrIlePheAlaSerCysAsp 80
Db |||||
QY 577 TCCAAACGGAGCTGATGTCACTGGCTCAATAGCTTTAGCAATTTGGATTTCTCTTGGCAATT 636
Db |||||
QY 81 SerLysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIle 100
QY 637 GGACATTTATTTGGCAATCAATTAATCTGTGTGCAGCATGAATCCCGCCGATCCTTTTGA 696
Db |||||
QY 101 GlyHisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGly 120
QY 697 CTGTGAGTTATCATGGAAATGGGAACCAATTCGATATATTTGGTGGGCCCATCATA 756
Db |||||
QY 757 GGAGCTGTCTCGCTGGTGGCTTTATGAGTATGTCTTCTGTCCAGATGTTGAAATTCAA 816
Db |||||
QY 141 GlyAlaValLeuAlaGlyAlaLeuTyrGluTyrValPheCysProAspValGluLeuLys 160
QY 817 CTGTGTTTAAAGAGCTTCAGCAAGCTGCCAGCAACAAAAGGAAGCTACATGGAG 876
Db |||||
QY 161 ArgArgLeuLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGlu 180
QY 877 GTGGAGGACACAGGAGTCAGGTAGAGACGATGACCTGATTTCTAAACCTCGAGTGGTG 936
Db |||||
QY 181 ValGluAspAsnArgSerGlnValGluThrGluAspLeuIleLeuLysProGlyValVal 200
QY 937 CATGTGATTCAGCTTACCGGGAGAGGAGAAAGGGGAAAGACCAATCTCGAGAGGTA 996
Db |||||
QY 997 TGTGCTTCAGTA 1008
Db |||||
QY 221 LeuSerSerVal 224

RESULT 11
Q6L7A0
ID Q6L7A0 PRELIMINARY; PRT; 218 AA.
AC Q6L7A0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Aquaporin type4 transcript variant c.
GN Name=AQP4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Ohya S., Inaiumi Y.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR EMBL; AB128929; BAD22823.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR GO; GO:0015288; F:porin activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000425; MIP.
DR Pfam; PF00230; MIP; 1.
DR PRINTS; PR01415; ANKYRIN.
DR PRINTS; PR00783; MINTRINSIC.
DR ProDom; PD000295; MIP; 1.
```

```
KW Porin; Transmembrane; Transport.
SQ SEQUENCE 218 AA; 23623 MW; AB8AB67A99D4504E CRC64;

Alignment Scores:
Pred. No.: 1,26e-73 Length: 218
Score: 1080.50 Matches: 218
Percent Similarity: 67.49% Conservative: 0
Best Local Similarity: 67.49% Mismatches: 0
Query Match: 49.54% Indels: 105
DB: 2 Gaps: 1
```

US-10-723-180-1 (1-1152) x Q6L7A0 (1-218)

```
QY 40 ATGAGTGCAGACCCACAGCAAGCGGTGGGTGAAGTGTGAGACCTTTGTGTACAGAGAG 99
Db 1 MetSerAspArgProThrAlaArgArgTrpGlyLysCysGlyProLeuCysThrArgGlu 20
QY 100 AACATCATGTGGTCTTCAAAGGGGTCTGGACTCAAGCTTCTCGAAAGCAGTCACAGCG 159
Db 21 Asn----- 21
QY 160 GAAATTTCTGGCCATGCTTATTTTGTCTCTCAGCCTGGGATCCACCATCAACTGGGGT 219
Db 21 ----- 21
QY 220 GGAACAGAAAAGCCCTTTACCGGTGCACATGTTCTCATCTCCCTTTGTTGGACTCAGC 279
Db 21 ----- 21
QY 280 ATTGCAACCATGTTGTCAGTGTCTTTGGCCATATCAGCGTGGCCACATCAACCTGCAGTG 339
Db 21 ----- 21
QY 340 ACTGTGGCCATGTTGTGCACCAAGAGATCAGCATCGCAAAGTCTGTCTTCTACATCGCA 399
Db 21 ----- 21
QY 400 GCCCAGTGCCTGGGGGCCATCATTTGAGAGCAGGAATCTCTATCTGTGTGCACACCTCCAGT 459
Db 22 -----IleIleGlyAlaGlyIleLeuTyrLeuValThrProSer 35
QY 460 GTGGTGGAGCCCTGGGAGTCACCATGTTTCATAGGAATCTTACCGCTGGTGCATGCTCTC 519
Db 36 ValValGlyGlyLeuGlyValThrMetValHisGlyAsnLeuThrAlaGlyHisGlyLeu 55
QY 520 CTGGTGTGAGTTGATAATCACATTTCAATGGTGTCTTACTATCTTTTGCAGCTGTGAATCC 579
Db 56 LeuValGluLeuIleThrPheGlnLeuValPheThrIlePheAlaSerCysAspSer 75
QY 580 AAACGGACTGATGTCACCTGGCTCAATAGCTTTAGCAATTTGATTTCTGTGCAATTGGA 639
Db 76 LysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyPheSerValAlaIleGly 95
QY 640 CATTTATTTGCAATCAATATATCTGTCGCCAGCATGAATCCGCCGATCTCTTTGGACT 699
Db 96 HisLeuPheAlaIleAsnTyrThrGlyAlaSerMetAsnProAlaArgSerPheGlyPro 115
QY 700 GCAGTTATCATGGGAATTTGGGAAACCATTTGGATATATTGGTGTGGCCCATCATAGGA 759
Db 116 AlaValIleMetGlyAsnTrpGluAsnHisTrpIleTyrTrpValGlyProIleIleGly 135
QY 760 GCTGTCTCGCTGGTGGCTTTATGAGTATGCTCTCTGTCCAGATGTTGAATTCAAAGT 819
Db 136 AlaValLeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluPheLysArg 155
QY 820 CGTTTAAAGAACCTTCAGCAAGCTGCCAGCAACAAAGGAAGCTACATGGAGGTG 879
Db 156 ArgPheLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMetGluVal 175
QY 880 GAGGACAAACAGGAGTCAGGTAGACCGGATCACTGATTTCTTAAACCTGGAGTGGTGCAT 939
Db 176 GluAspAsnArgSerGlnValGluThrAspAspLeuIleLeuLysProGlyValValHis 195
```

QY 940 GTGATTGACCTTGCACCGGAGAGGAGAAAGAGGGAAGACCAATCTGGAGAGGTATTG 999  
 Db |||||  
 196 ValIleAspValAspArgGlyGluGluLysGlyLysGlyLysAspGlnSerGlyGluValLeu 215  
 QY 1000 TCTTCAGTA 1008  
 Db |||||  
 216 SerSerVal 218

## RESULT 12

Q6AZD2 PRELIMINARY; PRT; 320 AA.  
 AC Q6AZD2;  
 DT 23-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Zgc:100922 protein.  
 GN Name=zgc:100922;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Singapore local strain; TISSUE=Embryo;  
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uslin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalusk D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=Singapore local strain; TISSUE=Embryo;  
 RA Director MGC Project;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -1- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.  
 DR EMBL; BC078213; AAH78213.1; -  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR000425; MIP.  
 DR Pfam; PF00230; MIP; 1.  
 DR PRINTS; PR00783; MINTRINSICP.  
 DR ProDom; PD000295; MIP; 1.  
 DR TIGRFAMs; TIGR00861; MIP; 1.  
 DR PROSITE; PS00221; MIP; 1.  
 KW Transmembrane; Transport.

SQ SEQUENCE 320 AA; 34427 MW; 6C9B4986CSA17C19 CRC64;  
 Alignment Scores:  
 Pred. No.: 2,44e-70 Length: 320  
 Score: 1037.50 Matches: 199  
 Percent Similarity: 77.71% Conservatives: 45  
 Best Local Similarity: 63.38% Mismatches: 63  
 Query Match: 47.57% Indels: 7  
 DB: 2 Gaps: 4

US-10-723-180-1 (1-1152) x Q6AZD2 (1-320)  
 QY 73 AAGTGTGGACCTTTGTGTACACAGAGAAC---ATCATGTGGCTTTCAAGGGGTCTGG 129  
 Db |||||  
 12 ArgCysValSerSerCysSerCysAsnAnSerIleMetAlaAlaPheLysGlyValTrp 31  
 QY 130 ACTCAAGCTTTCTGGAAAGCAGTCACAGCGGAATTTCTGGCCATGCTTATTTTGTCTC 189  
 Db |||||  
 32 ThrGlnGluPheTrpArgAlaValSerGlyGluPheLeuAlaMetIleIlePheValLeu 51  
 QY 190 CTCACCTGGGATCCACCATCAACTGGGGTGA---ACAGAAAAGCCTTTACCGGTGCAC 246  
 Db |||||  
 52 LeuSerLeuGlySerThrIleAsnTrpGlyAlaLysGlnGluAsnProProAlaAAsp 71  
 QY 247 ATGGTTCTCATCTCCCTTTGCTTGGACTCAGCATGCAACCATGGTGGAGTCTTGGC 306  
 Db |||||  
 72 LeuValLeuIleSerLeuCysPheGlyLeuSerIleAlaThrLeuValGlnCysPheGly 91  
 QY 307 CATATCAGCGGTGGCCACATCAACCTGCAGTGACTGTGGCCATGCTGTGGACCAAGAA 366  
 Db |||||  
 92 HisIleSerGlyAlaHisIleAsnProAlaValThrValAlaMetValAlaThrArgLys 111  
 QY 367 ATCAGCATGCGCAAGTCTGTCTTACATCGCAGCCAGTCTGGGGCCCATCATTTGA 426  
 Db |||||  
 112 LeuSerLeuAlaLysGlyValPheTyrluLeuAlaGlnCysLeuGlyAlaValValGly 131  
 QY 427 CGAGAAATCTCTATCTGTGTACACTCCAGTGTGGTGGAGGCTGGAGTCAACATG 486  
 Db |||||  
 132 AlaAlaIleLeuTyrluGlyValThrProAlaSerValArgGlyGlyMetGlyValThrSer 151  
 QY 487 GTTCATGAAATCTTACCGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 546  
 Db |||||  
 152 ValAsnGluGluIleSerAlaGlyHisAlaIleValIleGluLeuIleIleThrPheGlu 171  
 QY 547 TTGGTGTATTACTATCTTCCAGCTGTGATTCACAAAGGACTGTGTCTCTCTCTCTCT 606  
 Db |||||  
 172 LeuValPheThrValPheAlaThrCysAspProLysArgAsnAspLeuLysGlySerAla 191  
 QY 607 GCTTTAGCAATGGATTTCTGTTCATTTGACATTTTATTTGCAATCAATATATCTGT 666  
 Db |||||  
 192 AlaLeuAlaIleGlyLeuSerValCysIleGlyHisLeuPheAlaIleProTyrluGly 211  
 QY 667 GCACATGATATCCCGCCGATCTTTCAGCTGTGATTCATCGGAAATTTGGGAAAC 726  
 Db |||||  
 212 AlaSerMetAsnProAlaArgSerPheGlyProAlaValIleMetValLysTrpGlnAsp 231  
 QY 727 CATGTGATATATTGGTGTGGCCCATCATAGAGCTGCTCTCTCTCTCTCTCTCTCTAT 786  
 Db |||||  
 232 HisTrpValTyrluValGlyProLeuIleGlyGlyIleLeuAlaAlaValTyrluGlu 251  
 QY 787 TATGTCTTCTGTCCAGATGTGAATTAACCTCTCTTTTAAAGAGCCCTTCAGCAAGCT 846  
 Db |||||  
 252 TyrLeuPheCysProAspProAspLeuLysArgArgTyrluAlaAspValLeuSerLysSer 271  
 QY 847 GCCCAGCAACAAAGAGAGCTACATGAGGGTGGAGGACACAGAGTCAGGTAGAGCG 906  
 Db |||||  
 272 ProPheGlnMetGlu---ProTyrluValValAspThrAspSerTyrluProSerAspGln 290  
 QY 907 GATGACCTGTATCTTAAACCTGGAGTGTGTGATTTGATTTGATCGCGGGAGGAGGAG 966  
 Db |||||  
 291 AlaGlnLeuMetAlaLysGlnAlaAlaLeuArgValLeuAspLeu-----Glu 306  
 QY 967 AAGAGGGGAAAGACCAATCTGGAGAGGTATTTGCTTCTCAGTA 1008  
 Db |||||  
 307 LysLysGluArgGluSerThrGlyGluValLeuSerSerVal 320

## RESULT 13

O46426  
 ID O46426 PRELIMINARY; PRT; 178 AA.  
 AC O46426;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Aquaporin 4 (Fragment).  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=New Zealand;  
RA Carter E.P., Umenishi F., Matthey M.A., Verkman A.S.;  
RT "Increased water permeability across the blood-gas barrier in rabbit  
lungs in the first 24 hours after birth.";  
RL J. Clin. Invest. 0:0-0(1997).  
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
CC -|- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.  
DR EMBL; AF000312; AAB94409.1; -;  
DR HSSP; P29972; 1H61.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019867; C:outer membrane; IEA.  
DR GO; GO:0015288; F:porin activity; IEA.  
DR GO; GO:0015288; F:transporter activity; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR001110; ANK.  
DR InterPro; IPR000425; MIP.  
DR Pfam; PF00230; MIP; 1.  
DR PRINTS; PR01415; ANKYRIN.  
DR PRINTS; PR00783; MINTRINSCP.  
DR ProDom; PD000295; MIP; 1.  
KW Porin; Transmembrane; Transport.  
FT NON\_TER 1  
FT NON\_TER 178  
SQ SEQUENCE 178 AA; 19236 MW; CE43F9E18ACBAEA2 CRC64;  
  
Alignment Scores:  
Pred. No.: 1,09e-58 Length: 178  
Score: 884.00 Matches: 168  
Percent Similarity: 96.63% Conservatives: 4  
Best Local Similarity: 94.38% Mismatches: 6  
Query Match: 40.53% Indels: 0  
DB: 2 Gaps: 0  
  
US-10-723-180-1 (1-1152) x O46426 (1-178)  
  
QY 454 CCCAGTGGTGGGGGCTGGAGTCCACATGGTTCATGGAAATCTTACCCTGGTCAT 513  
Db 1 ProGlyValValGlyGlyLeuGlyValThrValHisGlyAsnLeuSerAlaGlyHis 20  
  
QY 514 GGTCTCCTGGTGGAGTTGATAATCACATTTCAATTTGGTGTCTTACTATCTTTGCCAGCTGT 573  
Db 21 GlyLeuLeuValGluLeuLeuLeuThrPheGlnLeuValPheThrIlePheProSerCys 40  
  
QY 574 GATTCACAAAGGAGCTGATCTACTGGCTCAATAGCTTTAGCAATTTGGATTTCTGTGCA 633  
Db 41 AspSerLysArgThrAspValThrGlySerIleAlaLeuAlaIleGlyIleSerValAla 60  
  
QY 634 ATTGGACATTTATTTGCAATCAATATATCTGTGCGCAGCATGAATCCGCCGATCCITTT 693  
Db 61 IleGlyHisLeuPheAlaIleAsnThrGlyAlaSerMetAsnProAlaArgSerPhe 80  
  
QY 694 GGACCTGCAAGTTATCATGGGAATTTGGGAAACCATTTGGATATATTGGGTGGGCCCATC 753  
Db 81 GlyProAlaValIleMetGlyAsnTrpGluAsnHisTrpIleTrpValGlyProIle 100  
  
QY 754 ATAGGAGCTGTCTCGCTGGTGGCCCTTTATGAGTATGCTTCTGTGTCAGATGTGAATTC 813  
Db 101 IleGlyAlaValLeuAlaGlyGlyLeuTyrGluTyrValPheCysProAspValGluLeu 120  
  
QY 814 AAACGTGTTTTAAGAGCCTTCAGAAAGCTGCCAGCAACCAACAAAGAGCTACATG 873  
Db 121 LysArgArgPheLysGluAlaPheSerLysAlaAlaGlnGlnThrLysGlySerTyrMet 140  
  
QY 874 GAGGTGGAGCAACAGGAGTCAAGTAGAGCGGATGACCTGATTTCTAAACCTGGAGTG 933  
Db 41 GluPheLeuAlaThrLeuIlePheValLeuLeuGlyValGlySerThrIleAsnTrpGly 60

Db 141 GluValGluAspAsnArgSerGlnValGluThrGluAspLeuIleLeuLysProGlyVal 160  
QY 934 GTGCATGTGATTGACCTGTGACCGGGAGAGAGAGAGAGAGAGAGAGAGAGCAATCT 987  
Db 161 MethisValIleAspIleAspArgGlyGluGlyLysGlyLysAspProSer 178  
  
RESULT 14  
Q8K4M1 PRELIMINARY; PRT; 149 AA.  
AC Q8K4M1;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Aquaporin-4 isoform M1 (Fragment).  
GN Name=Aqp4;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20419017; PubMed=10960499;  
RA Zelenin S., Gunmarson E., Alikina T., Bondar A., Aperia A.;  
RT "Identification of a new form of AQP4 mRNA that is developmentally  
expressed in mouse brain.";  
RL Pediatr. Res. 48:335-339(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Bondar A., Alikina T., Zelenin S.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Alikina T.Y., Zelenin S.M., Bondar A.A.;  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
CC -|- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.  
DR EMBL; AP219992; AAM81576.1; -;  
DR HSSP; P29972; 1H61.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019867; C:outer membrane; IEA.  
DR GO; GO:0015288; F:porin activity; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000425; MIP.  
DR Pfam; PF00230; MIP; 1.  
DR PRINTS; PR00783; MINTRINSCP.  
DR ProDom; PD000295; MIP; 1.  
DR PROSITE; PS00221; MIP; 1.  
KW Porin; Transmembrane; Transport.  
FT NON\_TER 149  
FT NON\_TER 149  
SQ SEQUENCE 149 AA; 15663 MW; BA0B8F89359B708C CRC64;  
  
Alignment Scores:  
Pred. No.: 3.08e-45 Length: 149  
Score: 707.00 Matches: 135  
Percent Similarity: 94.59% Conservatives: 5  
Best Local Similarity: 91.22% Mismatches: 8  
Query Match: 32.42% Indels: 0  
DB: 2 Gaps: 0  
  
US-10-723-180-1 (1-1152) x Q8K4M1 (1-149)  
  
QY 40 ATGAGTGCAGACCCACAGCGGCTGGGGTGAAGTGTGGACCTTTGTGTACGAGAG 99  
Db 1 MetSerAspArgAlaAlaAargArgTrpGlyLysCysGlyHisSerCysSerArgGlu 20  
  
QY 100 AACATCATGTGGCTTTCAAAGGGCTCTGGAAGCTTCTGAGAAAGAGTGCACGCG 159  
Db 21 SerIleMetValAlaPheLysGlyValTrpGlnAlaPheTrpLysAlaValSerAla 40  
  
QY 160 GAATTTCTGGCCATGCTTATTATTTTCTCTCCAGCTGGGATCCACCATCAACTGGGT 219  
Db 41 GluPheLeuAlaThrLeuIlePheValLeuLeuGlyValGlySerThrIleAsnTrpGly 60

QY 220 GGAACAGAAAAGCTTTACCGGTGACATGGTTCTCATCTCCCTTGTGACTCAGC 279  
 |||:|||||  
 Db 61 GlySerGluAsnProLeuValAspMetValLeuIleSerLeuCysPheGlyLeuSer 80  
 |||:|||||  
 QY 280 ATTGCACCATGGTGCAGCTGCTTTGGCCATATCAGCGGTGGCCACATCAACCTGCGAGTG 339  
 |||:|||||  
 Db 81 IleAlaThrMetValGlnCysPheGlyHisIleSerGlyGlyHisIleAsnProAlaVal 100  
 |||:|||||  
 QY 340 ACTGTGGCCATGGTGTGCACCAAGAGATCAGCATCGCCAAAGTCTGTCTTACATCGCA 399  
 |||:|||||  
 Db 101 ThrValAlaMetValCysThrArgLysIleSerIleAlaLysSerValPheIleIle 120  
 |||:|||||  
 QY 400 GCCAGTGCCTGGGGCCCATCATTTGGAGCAGGAATCCTCTATCTGTGCACATCCCACT 459  
 |||:|||||  
 Db 121 AlaGlnCysLeuGlyAlaIleIleGlyAlaGlyIleLeuTyLeuValThrProSer 140  
 |||:|||||  
 QY 460 GTGTGGGAGGCTGGGAGTCACC 483  
 |||:|||||  
 Db 141 ValValGlyGlyLeuGlyValThr 148  
 |||:|||||

RESULT 15  
 AQP\_AEDAE  
 ID AQP\_AEDAE STANDARD; PRT; 249 AA.  
 AC Q9NH7;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Aquaporin AQPRe.a.  
 OS Aedes aegypti (Yellowfever mosquito).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.  
 ON NCBI\_TaxID=7159;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UTMB; TISSUE=Malpighian tubules;  
 RX MEDLINE=20428297; PubMed=10971718;  
 RA Pietrantonio P.V., Jagge C., Keeley L.L., Ross L.S.;  
 RT "Cloning of an aquaporin-like cDNA and in situ hybridization in adults  
 of the mosquito Aedes aegypti (Diptera: Culicidae).";  
 RL Insect Mol. Biol. 9:407-418(2000).  
 CC -!- FUNCTION: Forms a water-specific channel (By similarity).  
 CC -!- SUBUNIT: Homotetramer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: Localized in tracheolar cells associated with  
 the Malpighian tubules.  
 CC -!- DOMAIN: Aquaporins contain two tandem repeats each containing  
 three membrane-spanning domains and a pore-forming loop with the  
 signature motif Asn-Pro-Ala (NPA).  
 CC -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF218314; AAF64037.1; -.  
 CC HSSP; P29972; 1H61.  
 CC InterPro; IPR000425; MIP.  
 CC Pfam; PF00230; MIP; 1.  
 CC PRINTS; PR00783; MINTREINSCP.  
 CC ProDom; PD000295; MIP; 1.  
 CC TIGRFAMs; TIGR00861; MIP; 1.  
 CC PROSITE; PS00221; MIP; 1.  
 CC Repeat; Transmembrane; Transport.  
 KW DOMAIN 1 26 Cytoplasmic (Potential).  
 FT TRANSMEM 27 47 1 (Potential).  
 FT DOMAIN 48 57 Extracellular (Potential).  
 FT TRANSMEM 58 78 2 (Potential).  
 FT DOMAIN 79 100 Cytoplasmic (Potential).

FT TRANSMEM 101 121 3 (Potential).  
 FT DOMAIN 122 138 Extracellular (Potential).  
 FT TRANSMEM 139 159 4 (Potential).  
 FT DOMAIN 160 169 Cytoplasmic (Potential).  
 FT TRANSMEM 170 190 5 (Potential).  
 FT DOMAIN 191 211 Extracellular (Potential).  
 FT TRANSMEM 212 232 6 (Potential).  
 FT DOMAIN 233 249 Cytoplasmic (Potential).  
 FT SITE 82 84 NPA 1.  
 FT SITE 198 200 NPA 2.  
 SQ SEQUENCE 249 AA; 26111 MW; F8584962C5F169FE CRC64;

## Alignment Scores:

Pred. No.: 2,246-41 Length: 249  
 Score: 656.50 Matches: 118  
 Percent Similarity: 69.72% Conservative: 34  
 Best Local Similarity: 54.13% Mismatches: 63  
 Query Match: 30.10% Indels: 3  
 DB: 1 Gaps: 1

US-10-723-180-1 (1-1152) x AQP\_AEDAE (1-249)

QY 142 TGGAAAGCAGTCACAGCGGAATTTCTGCCCATGCTTATTTTGTCTCTCTCAGCTCGGA 201  
 |||:|||||  
 Db 23 TrpArgMetLeuValAlaGluPheLeuGlyThrPhePheLeuValSerIleGlyIleGly 42  
 |||:|||||  
 QY 202 TCCACCATCAATCGGGTGGAAACAGAAAGCCTTTACCGGTGCACATGTTCTCATCTCC 261  
 |||:|||||  
 Db 43 SerThrMetGlyTrpGlyGlyAspTyrAlaPro-----ThrMetThrGlnIleAla 59  
 |||:|||||  
 QY 262 CTTTGTCTTGGACTCAGCATGCAACCATGTTGGAGTGTGTCACAGGAAGATCAGCATCGCG 321  
 |||:|||||  
 Db 60 PheThrPheGlyLeuValValAlaThrLeuAlaGlnAlaPheGlyHisValSerGlyCys 79  
 |||:|||||  
 QY 322 CACATCAACCTCGCAGTGCATGTGCCCATGTTGGAGTGTGTCACAGGAAGATCAGCATCGCG 381  
 |||:|||||  
 Db 80 HisIleAsnProAlaValThrIleGlyLeuMetIleThrAlaAspIleSerIleLeuLys 99  
 |||:|||||  
 QY 382 TCTGTCTTCTACATCGACCCAGTGTGGGGCCCATCATTTGGAGCAGGAATCCTCTAT 441  
 |||:|||||  
 Db 100 GlyAlaPheTyrIleValSerGlnCysValGlyAlaIleAlaGlyAlaLeuIleLys 119  
 |||:|||||  
 QY 442 CTGGTTCACACCTCCAGTGTGGTGGAGCCCTGGGAGTCACCATGGTTTCATGGAATCTT 501  
 |||:|||||  
 Db 120 AlaAlaThrProSerAspValIleGlyGlyLeuGlyValThrGlyIleAspProArgLeu 139  
 |||:|||||  
 QY 502 ACCGCTGTCTATGTTCTCTCTGTGTGAGTTGATATCATCATTTCAATTTGGTGTGTATATC 561  
 |||:|||||  
 Db 140 ThrAlaGlyGlnGlyValMetMetGluAlaLeuIleThrPheIleLeuValPheValVal 159  
 |||:|||||  
 QY 562 TTTGCCAGCTGTGATTCACAAACGGACTCATGTCCTGCTCAATAGCTTTAGCAATTTGGA 621  
 |||:|||||  
 Db 160 HisGlyValCysAspAsnArgSerAspIleLysGlySerAlaProLeuAlaIleGly 179  
 |||:|||||  
 QY 622 TTTTCTGTGCAATTTGGACATTTATTTGCAATATATCTGTGTGCGCCAGCATGAATCCC 681  
 |||:|||||  
 Db 180 LeuSerIleThrAlaGlyHisLeuSerAlaIleLysTyrThrGlyAlaSerMetAsnPro 199  
 |||:|||||  
 QY 682 GCGCGATCCTTTGGACCTGCAGTTATCATCGGAATTTGGAAACCATTTGGATATATTTGG 741  
 |||:|||||  
 Db 200 AlaArgSerPheGlyProAlaValValMetGlyAsnTrpThrAspGlnTrpValTyrTrp 219  
 |||:|||||  
 QY 742 GTTGGGCGCCATCAGGAGCTGTCTCGCTGGTGGCTTTTATGATGATGTCTTCTTC 795  
 |||:|||||  
 Db 220 ValGlyProIleValGlyGlyIleLeuAlaGlyAlaValTyrArgLeuPhePhe 237  
 |||:|||||

Search completed: September 8, 2005, 12:13:42  
 Job time : 238 secs

**This Page Blank (uspto)**